14.771 Development Economics: Microeconomic Issues and Policy Models
Fall 2008

Why study (micro) development economics ?

(and what is there to study?)

Astonishing differences between countries

- Richest country in the World…
- Poorest country in the World…
- At PPP?
- Translate into astonishing differences in well being (e.g. infant mortality)—see tables.
- Lucas's well known phrase is certainly justified: once you start thinking about it, there is no reason to stop!
- This is the object of development economics: why do poor countries stay poor? why do poor people in poor countries have such short and hard lives?
- And, at least in our view, what can be done about it (a resolutely normative view of economics)?

Perhaps we just need to study growth (macroeconomics)?

- Banerjee (2008) “Big answers for Big questions”
- Will the tide lift all boats (growth is sufficient for poverty reduction)?
  - No correlation between growth and increase in inequality (Sala-i-Martin; Dollar and Kray): growth is correlated about one for one with poverty reduction
  - Widely debated (Data is poor; Debate on counting the poor is hot and raging—e.g. India. Experiences vary: the effect of growth on poverty is different from Country to Country -Ravallion)
- Even if true, it could still be the case that poverty reduction cause growth, rather than the other way around…

Do we know what makes growth happen?

- Aggregate production function:

$$Y = f(A, K, H)$$

- Mankiw, Romer, Weil (1992) find that 80% of the difference in growth across countries can be explained by savings rate and investment in human capital.
- Caselli (2005): Two thirds of the variation remains unexplained by these factors.
- Problem of attributing causality: Does education cause growth of the other way around (Bills and Klenow).

What do we know about the effect of Policies on Growth?

- Acemoglu, Johnson, Robinson (2001, 2002); La Porta, Lopez de Silanes, Shleifer, Vishny (many papers): Very long run factors matter: who colonized you? did they decided to stay?
- Once you account for that, short run policies do not appear to be very correlated with growth (Rodrik, Subramanian, Trebbi, 2002).
- Very different policies lead to very different outcomes
- It is difficult to learn from country experience in a statistical sense, because each is so specific (just one China).
- In contrast, we can learn a lot about the effect of a specific policy on a specific outcome in a particular place by looking at micro- data from this country (we will do a lot of this this semester).

Can we describe the economy with an aggregate production function?

- Aggregate production function:

$$Y = f(A, K, H)$$

- We don’t think of the economy as one big machine…
- What justifies the aggregate production function is the assumption that each factor (human capital, capital) gets allocated to its most productive use: the marginal return of every investment is equalized.
- This is a reasonable assumption, since this is what market forces should ensure.
- And in this case, it does not matter where exactly capital and human capital are invested: we can think of one aggregate stock (Solow).

Can we describe the economy with an aggregate production function?

- One nice thing about this assumption is that it can be tested: is it true that the returns to capital are equalized within an economy: We can look at the returns to capital in various places
- And the answer it… it is not at all true!! Returns to capital (and human capital) vary much more within countries than across countries, which is the puzzle that set Lucas on his pathbreaking effort to think about growth.
- What are the barriers that would prevent the returns to capital to equalize in developing countries?
- We better start thinking about that in some details…
- We won't get a good understanding of growth until we understand the micro- mechanisms at play: once we do, we will be able to think better about how everything aggregates up.

What is there to study? A birdview of the course…

The economic lives of the poor- What is middle class about the middle classes? (Banerjee-Duflo, 2007,2008)

- Bring together data on the **very poor**, defined as those with consumption below $1.08 a day at 1993 PPP.
- And the **poor** defined as those below $2.16 a day at 1993 PPP
- And the middle classes (2-4 and 6-10)
- From various detailed household surveys
- Gives us a sense of how the poor live, what are the big questions to be asked.

Family Sizes

A young population…

- Look at table 2
  - Large families
  - Many children
  - Big changes from the very poor to the less poor
  - Do large family sizes cause poverty or the other way around?

How do the poor spend their money?

How the poor spend their money

As a Share of Total Consumption

	Food	Alcohol/ Tobacco	Education	Health
Rural				
Cote d'Ivoire	64.4%	2.7%	5.8%	2.2%
Guatemala	65.9%	0.4%	0.1%	0.3%
India - Udaipur	56.0%	5.0%	1.6%	5.1%
India - UP/Bihar	80.1%	3.1%	0.3%	5.2%
Indonesia	66.1%	6.0%	6.3%	1.3%
Mexico	49.6%	8.1%	6.9%	0.0%
Nicaragua	57.3%	0.1%	2.3%	4.1%
Pakistan	67.3%	3.1%	3.4%	3.4%
Panama	67.8%		2.5%	4.0%
Papua New Guinea	78.2%	4.1%	1.8%	0.3%
Peru	71.8%	1.0%	1.9%	0.4%
South Africa	71.5%	2.5%	0.8%	0.0%
Timor Leste	76.5%	0.0%	0.8%	0.9%

How the poor spend their money

As a Share of Total Consumption

Rural	Entertainment	Festivals	% HHs with any Festival Expenditure
Cote d'Ivoire	0.0%	1.3%	59.9%
Guatemala		0.1%	7.7%
India - Udaipur	0.0%	14.1%	99.4%
India - UP/Biha	0.1%	2.2%	
Indonesia	0.0%	2.2%	80.3%
Mexico	0.7%	0.0%	2.7%
Nicaragua	0.0%	0.0%	1.8%
Pakistan	0.3%	2.4%	64.8%
Panama	0.6%	0.0%	0.0%
Papua New Gui	0.2%	1.5%	21.7%
Peru	0.0%		
South Africa	0.1%	3.2%	90.3%
Timor Leste	0.0%	0.0%	49.0%

Food?

- Are the very poor spending every marginal penny on they can on getting more food?
- The share of expenditure between 78% in Papua New Guinea and 50% in Mexico.
- Other large items include:
  - Tobacco/alcohol (up to 8%)
  - Festivals (up to 14% when asked in detail)
- The poor spend 30% of their food budget on rice and wheat which cost between 70% and 100% more per calorie than other grains
- Food share is falling over time in most countries. So, more surprisingly, is calory share…
- Yet, Average BMI is 17.8 in Udaipur (a poor Indian district) compared to a cut-off for under-nourished of 18.5 & 55% are anemic.

- Are the poor eating enough?

How the poor spend their money

As a Share of Total Consumption

	Food	Alcohol/ Tobacco	Education	Health
Rural				
Cote d'Ivoire	64.4%	2.7%	5.8%	2.2%
Guatemala	65.9%	0.4%	0.1%	0.3%
India - Udaipur	56.0%	5.0%	1.6%	5.1%
India - UP/Bihar	80.1%	3.1%	0.3%	5.2%
Indonesia	66.1%	6.0%	6.3%	1.3%
Mexico	49.6%	8.1%	6.9%	0.0%
Nicaragua	57.3%	0.1%	2.3%	4.1%
Pakistan	67.3%	3.1%	3.4%	3.4%
Panama	67.8%		2.5%	4.0%
Papua New Guinea	78.2%	4.1%	1.8%	0.3%
Peru	71.8%	1.0%	1.9%	0.4%
South Africa	71.5%	2.5%	0.8%	0.0%
Timor Leste	76.5%	0.0%	0.8%	0.9%

Health in the Household

	In Last Month				
	Percent of HH Members Sick	A Household's Average # of Consultations	Percent of Households that met At Least Once with a Consultant: Public	Private	Infant Mortality
Rural					
Cote d'Ivoire	21.4%	1.28	49.7%	3.2%	6.2%
Guatemala					6.2%
India - Hyderabad					
India - Udaipur	46.1%	0.11	20.1%	58.1%	10.0%
India - UP/Bihar	12.5%	0.81	13.9%	47.3%	7.7%
Indonesia	24.2%	0.77	20.7%	27.3%	3.4%
Mexico	46.3%	1.11	47.7%	0.0%	6.9%
Nicaragua	34.9%	0.15	46.0%	5.0%	
Pakistan	28.0%	0.45	24.0%	48.8%	16.7%
Panama	15.2%	0.10	23.8%	0.0%	
Papua New Guinea					
Peru	11.1%	0.10	20.9%	8.5%	
South Africa	12.5%	0.12	16.4%	6.9%	8.6%
Tanzania	13.2%	0.07	23.2%	14.0%	8.7%
Timor Leste	11.7%	0.21	30.2%	0.5%	

Education

	Percent of Children in School			
	Female, Age:		Male, Age:	
	7-12	13-18	7-12	13-18
Rural				
Cote d'Ivoire	32.3%	22.8%	45.5%	21.1%
India - Udaipur	60.7%	13.0%	82.6%	24.7%
India - UP/Bihar	51.4%	20.2%	72.1%	51.2%
Indonesia	93.4%	45.9%	82.4%	39.3%
Mexico	94.5%	56.5%	93.5%	38.6%
Nicaragua	67.5%	38.0%	65.4%	27.5%
Pakistan	30.7%	9.2%	64.1%	41.3%
Panama	79.0%	14.6%	85.1%	27.0%
Papua New Guinea	53.0%	33.5%	71.4%	70.9%
Peru	94.2%	64.7%	93.3%	73.7%
South Africa	83.6%	87.5%	80.5%	76.9%
Tanzania	51.2%	53.3%	47.2%	61.4%
Timor Leste	76.6%	89.7%	80.0%	86.8%

Health and Education

- Lots of illnesses
- Lots of visits, to private and public doctors
- High expenditures on health, which increase steeply with income
- Most children are in school, even among the poor… But expenditure on education increases steeply with income as well.

Saving and Accumulation

What do the poor own

Percent of Households with:

	Radio	Television	Bicycle	Land
Rural				
Cote d'Ivoire	43.3%	14.3%	34.4%	62.7%
Guatemala	58.5%	20.3%	23.1%	36.7%
India - Hyderabad				
India - Udaipur	11.4%	0.0%	13.5%	98.9%
India - UP/Bihar	28.3%	7.3%	65.8%	
Indonesia		26.5%		49.6%
Mexico			41.3%	4.0%
Nicaragua	59.3%	8.3%	11.1%	50.4%
Pakistan	23.1%		27.0%	30.4%
Panama	43.6%	3.3%	0.0%	85.1%
Papua New Guinea	18.0%	0.0%	5.3%	
Peru	73.3%	9.8%	9.8%	65.5%
South Africa	72.2%	7.2%	20.0%	1.4%
Tanzania		0.0%		92.3%
Timor Leste	14.3%	0.6%	0.9%	95.2%

Ownership

- In the median country a majority of the rural poor own land.
- Other than that they own very little: in the Udaipur sample, 10% or less have a chair or a table
- In the median country less than 15% have a bicycle and less than 10% own a television.
- In Udaipur very few possible business assets: Less than 1% own a tractor, a bullock cart, a motorized cycle, a fan

Savings and fluctuations

- They seem to have very little by way of savings
- Few bank Accounts
- Large fluctuations: Udaipur
  - in the last year 37% of the very poor households the adults starved for an entire day.
  - 45% cut the size of their meals
  - Only 57% report that they have enough to eat throughout the year
  - Those families that report missing meals are 0.23 standard deviations less well-being

Questions

- Why so many children?
- Why don’t they eat more?
- Why so many illnesses?
- Why so many private visits?
- Why so few assets?
- Are the poor getting enough education?

How do the poor earn their money?

How the poor earn their money: Occupation

Rural	Percent of Households that own land	Median Ares Of Land Owned	Percent of Households in which At Least One Member: Is Self Employed In Agriculture	Other
Cote d'Ivoire	62.7%	300	37.2%	25.9%
Guatemala	36.7%	29	64.4%	22.6%
India - Udaipur	98.9%	60	98.4%	5.9%
India - UP/Bihar		40	72.1%	40.2%
Indonesia	49.6%	60	49.8%	36.6%
Mexico	4.0%		4.9%	20.4%
Nicaragua	50.4%	280	54.7%	11.6%
Pakistan	30.4%	162	72.1%	35.5%
Panama	85.1%	300	69.1%	17.7%
Peru	65.5%	150	71.7%	25.2%
South Africa	1.4%		0.0%	9.1%
Tanzania	92.3%	182		
Timor Leste	95.2%	100	78.5%	12.0%

How the poor earn their money: Occupation

	Percent of Households in which At Least One Member: Works for a Wage or Salary in		Percent of HHs That Receive Income From Multiple Sectors
	Agriculture	Other	
Rural			
Cote d'Ivoire	52.4%	78.3%	72.1%
Guatemala	31.4%	86.4%	83.8%
India - Udaipur	8.5%	90.7%	94.0%
India - UP/Bihar	2.0%	18.9%	41.8%
Indonesia	31.1%	34.3%	50.4%
Mexico	2.8%	72.6%	13.2%
Nicaragua	0.3%	42.8%	18.4%
Pakistan	32.6%	50.8%	66.8%
Panama	0.0%	0.0%	19.2%
Peru			34.8%
South Africa	27.9%	26.6%	0.4%
Tanzania			
Timor Leste			10.4%

Little bit of everything

- In most countries a majority of household get income from self-employment
- Agriculture is far from being everything! Even in rural areas self-employment is both in agriculture and outside it and many do both)
- Compared to the middle class, the poor are as likely to be entrepreneurs
- The next commonest occupation is wage work, including a lot of casual work. Many households do both.
- The big difference between the poor and the middle classes is the stability of the job (casual labor vs salaried job)
- 20% of the households in rural Udaipur say agriculture is their main source of earnings. 75% say their main earnings come from wage work. Yet almost all of them own land and cultivate it.

Non-Agriculture Enterprises Owned by Household

	Percent of HHs with at least One Non-Agricultural Business	In Each Business: Average Number of Employees		Percent of Businesses that Own:	
		Paid Workers	Paid + Unpaid	Vehicles	Machines
Rural					
Cote d'Ivoire	66.4%	0.14	2.48	2.6%	66.5%
India - UP/Bihar	35.0%				
Indonesia	29.4%	0.11	1.55	0.0%	
Mexico	7.8%	0.59	2.16	0.0%	
Nicaragua	14.0%	0.08	1.39	7.5%	0.0%
Pakistan	34.3%	0.13	1.16	36.7%	0.0%
Panama	15.2%	0.00	1.58	0.0%	
Papua New Guinea					
Peru	34.5%		1.50		

The businesses the poor run

- Families often run multiple businesses
- Businesses are very small. Almost no outside labor
- In Hyderabad only 20% of businesses have their own premises and the commonest businesses assets are pushcarts, scales and tables.
- In Hyderabad the main businesses are tailoring, fruit/vegetable selling, general stores, telephone booths, selling milk, driving a small taxi: Skills?

Migration

	Percent of Adults who Have Migrated	
	Since Birth	For Work
Rural		
Cote d'Ivoire	26.9%	11.1%
Guatemala		
Indonesia	34.3%	30.8%
Mexico	48.7%	45.6%
Nicaragua	22.4%	5.6%
Pakistan	16.7%	3.7%
Panama	34.8%	0.6%
Papua New Guinea	4.8%	
Peru	15.3%	6.7%
Timor Leste	61.6%	

Migration

- In most countries in our sample long-term migration for work is rare among the poor
- More frequent among the rich
- Temporary migration on the other hand seems quite common. In 60% of the very poor households in Udaipur someone had migrated for work.
- Average duration of a completed episode is 40 days.
- Total time spent away in a year on average is 18 weeks.
- Yet for many of them it provides the majority of their income.

Questions

- Why are the businesses so small?
- Why so many entrepreneurs?
- Why don’t they migrate for longer?
- Why so little specialization?

The Economic environment of the poor

Market for Credit and Savings and the poor

	Savings Group	Shopkeeper	Villager	Relative	Friend	Other	% HH with a Savings Account
Rural							
Cote d'Ivoire	0.0%		94.3%		0.0%	0.0%	79.5%
India - Udaipur	2.6%	36.4%	4.0%	21.6%	2.1%	2.8%	6.4%
India - UP/Bihar	1.5%		60.9%		0.0%	1.3%	
Indonesia	17.8%	0.0%	0.0%		0.0%	51.3%	6.6%
Mexico				53.5%	18.3%	8.3%	6.2%
Nicaragua							
Pakistan	0.0%	15.8%	11.2%	38.1%	29.0%	3.7%	11.7%
Panama							0.5%
Papua New Guinea							
Peru		9.2%		23.9%			0.5%
South Africa		71.3%			26.1%	16.7%	
Timor Leste							13.4%

Credit and savings

- An enormous fraction of the poor in many countries carry some debt
- Almost none of them got it from the formal banking sector
- High interest rate: often more than 3% a month!
- Very few people have formal savings accounts.

Market for Insurance and the poor

	Percent of Total Households with Insurance:		
	Any Type	Health	Life
Rural			
Cote d'Ivoire			
Guatemala			
India - Hyderabad			
India - Udaipur			3.8%
India - UP/Bihar	9.2%	4.7%	3.8%
Indonesia	6.0%	3.9%	0.0%
Mexico		50.7%	
Nicaragua	0.0%	5.5%	
Pakistan			
Panama		0.0%	0.0%
Papua New Guinea			
Peru		5.6%	0.0%
South Africa	5.4%		
Tanzania			
Timor Leste			

Insurance

- Very little formal insurance against anything
- Large fluctuation in food consumption/school enrollment in response to exogenous income shocks

Land

- The very poor are very likely to have land
- The less poor some times even more likely, and they often even hold less!

Questions

- Why no formal credit and insurance?
- Are informal systems good substitute for formal systems?
- Why are interest rates so high?
- How can the poor afford such high interest rates?
- Why don't they repay their debt as a way to save?
- Why do so many people have land?

Public goods and infrastructures

Economics environment of the poor: Basic infrastructure

Percent of Households with:

	In-House Tap Water	Toilet/ Latrine	Electricity
Rural			
Cote d'Ivoire	11.8%	27.1%	45.1%
Guatemala	37.7%	50.5%	29.9%
India - Udaipur	0.0%	0.0%	8.3%
India - UP/Bihar	1.9%	3.4%	8.7%
Indonesia	5.6%	30.5%	96.9%
Mexico			99.0%
Nicaragua	12.3%	59.0%	16.4%
Pakistan	9.9%	28.5%	55.5%
Panama		37.7%	0.0%
Papua New Guinea	1.7%	95.2%	2.0%
Peru	29.7%		12.2%
South Africa	4.4%	58.9%	5.6%
Tanzania	0.7%	91.6%	1.1%
Timor Leste	2.3%	31.3%	8.8%

Poor quality of public good and infrastructure

- Enormous variation across country/income groups in availability of physical infrastructure (water/electricity)
- Health and education infrastructure
  - Expenditures of the poor on education are not too high despite fairly high enrollment
  - However… private school enrollment rises steeply with income
  - So does expenditure per health visits
- Quality of public infrastructure ??

The game plan…

This semester is about building blocks

- Household behavior (health, education, techology adoption)
- Public infrastructure (service delivery; infrastructure: dams/phone/electricity etc)
- Markets (land; credit; savings; product)

And what comes next…

- Macro: incorporating the building blocks in good models (14.772; Macro 1)
- Political economy: what are the determinants of the delivery of these public goods (14.773?)

14.771 Development Economics: Microeconomic Issues and Policy Models
Fall 2008

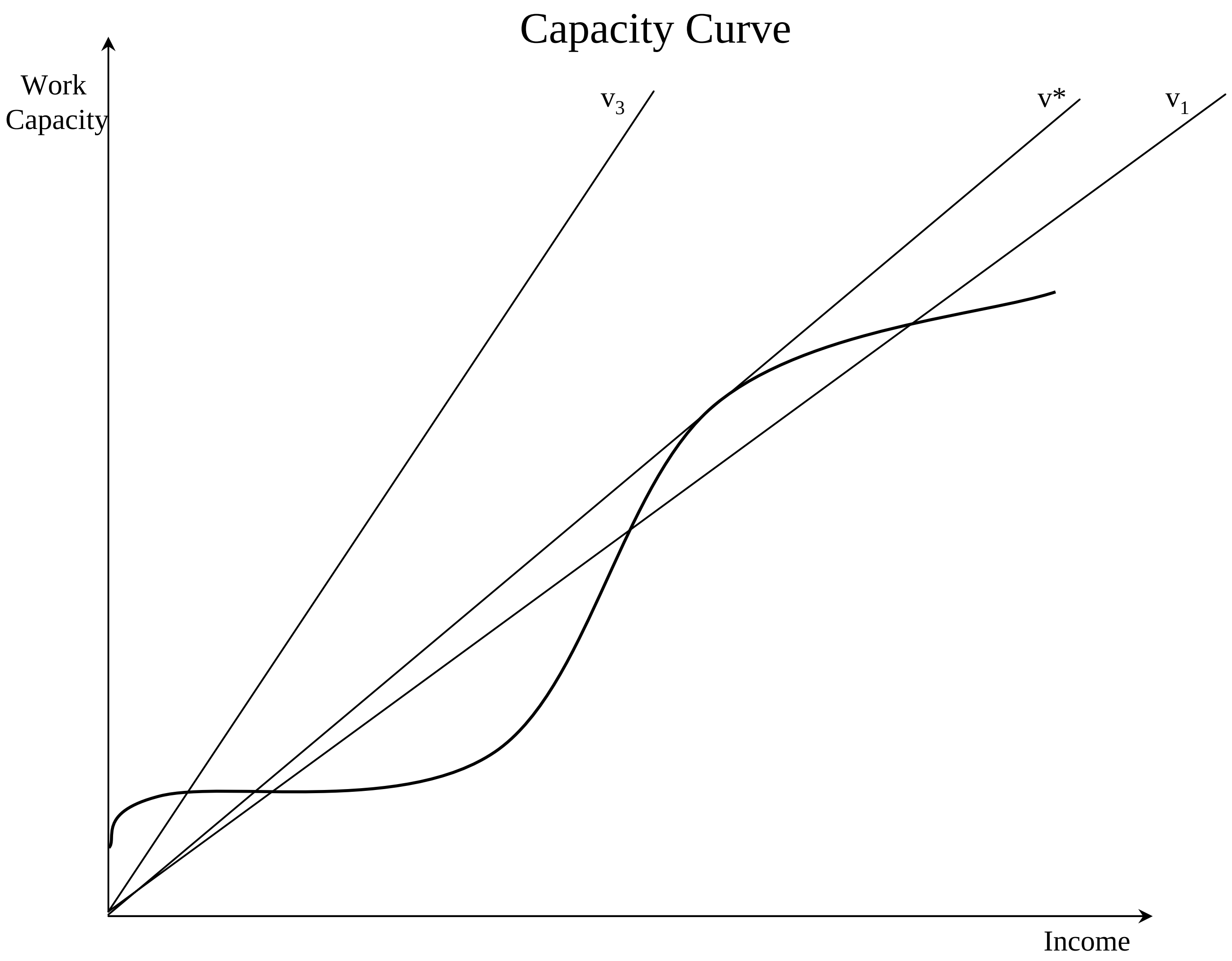


Figure 1: The Capacity Curve
The Piece Rate

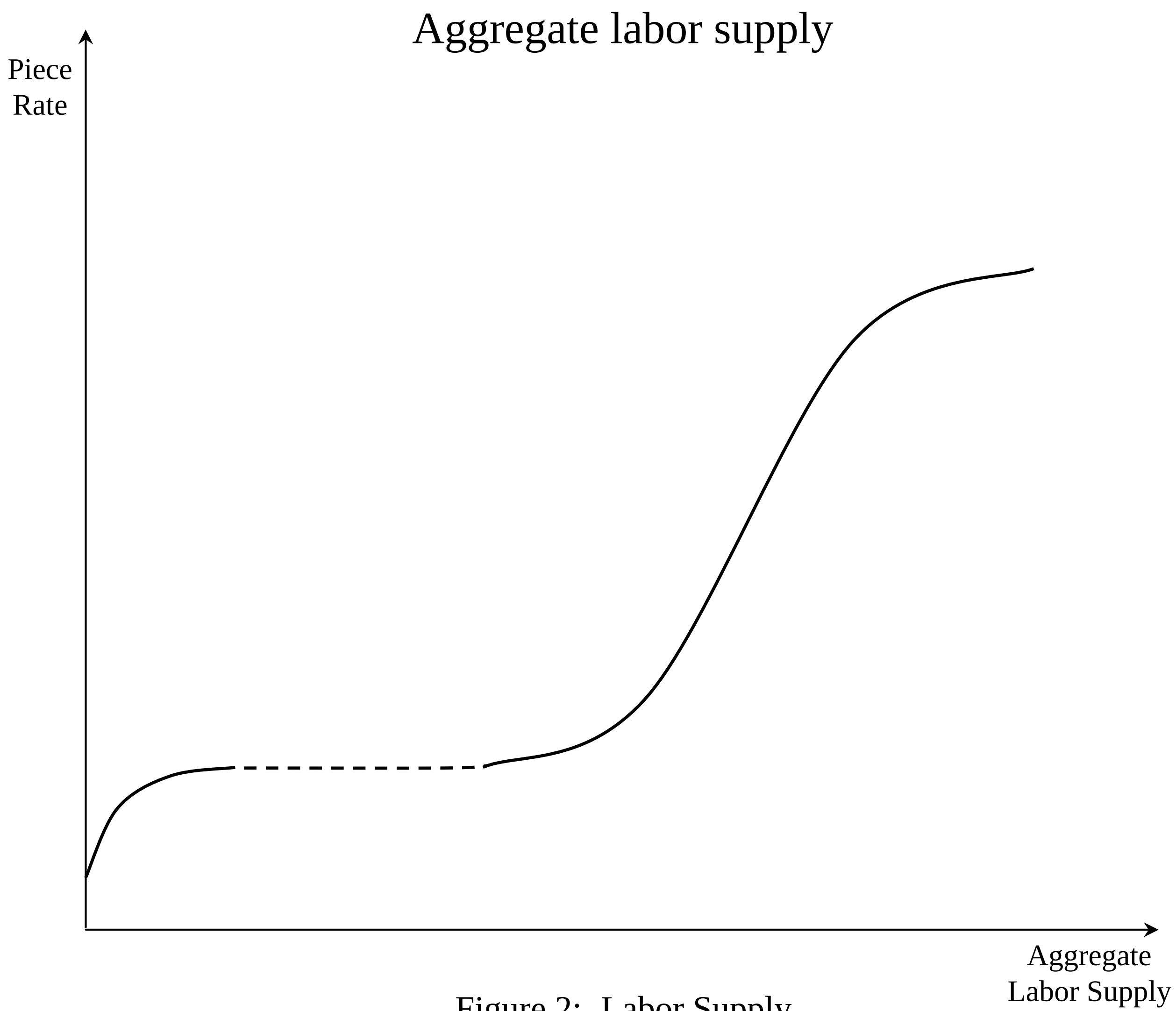


Figure 2: Labor Supply

Possible equilibria

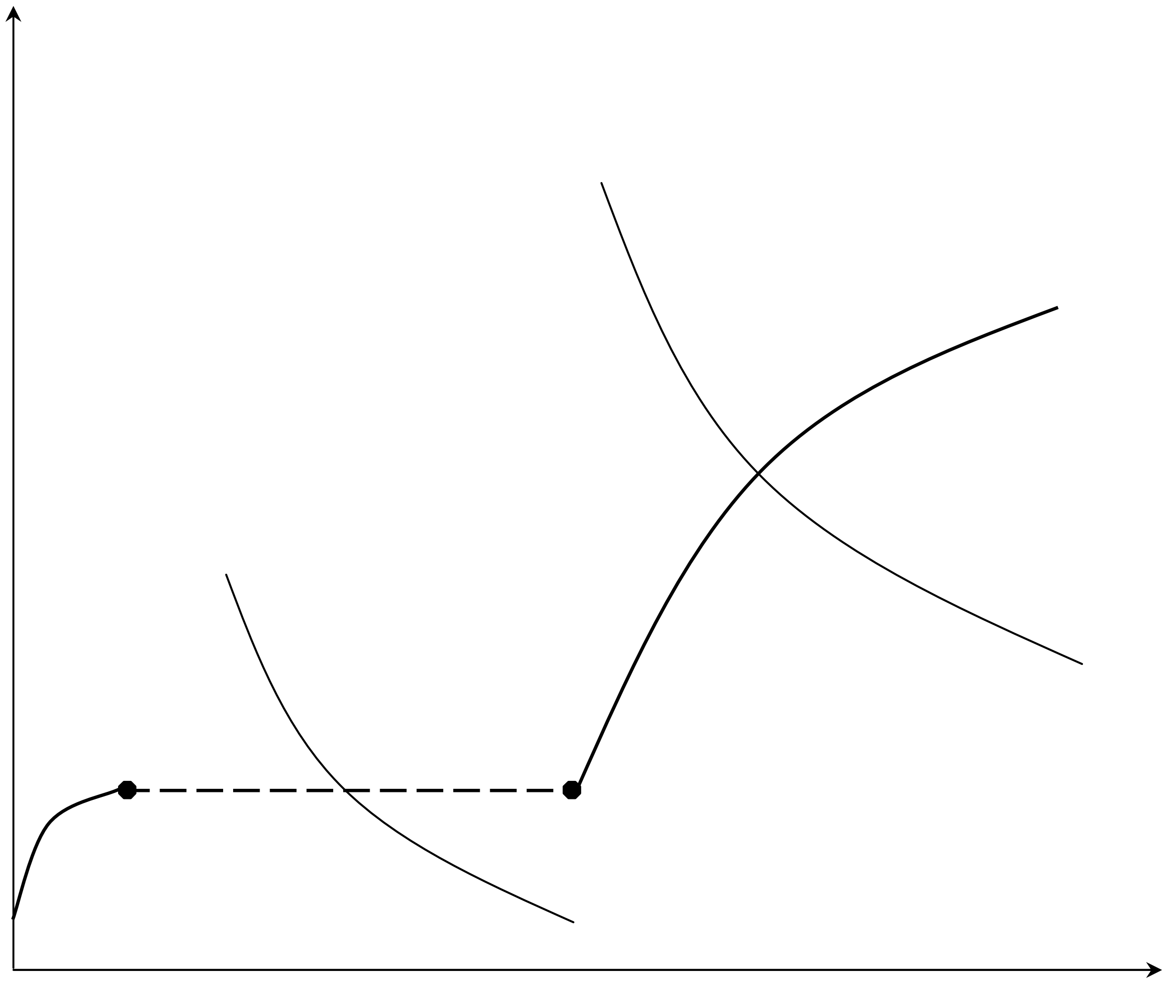


Figure 3: Possible Equilibria

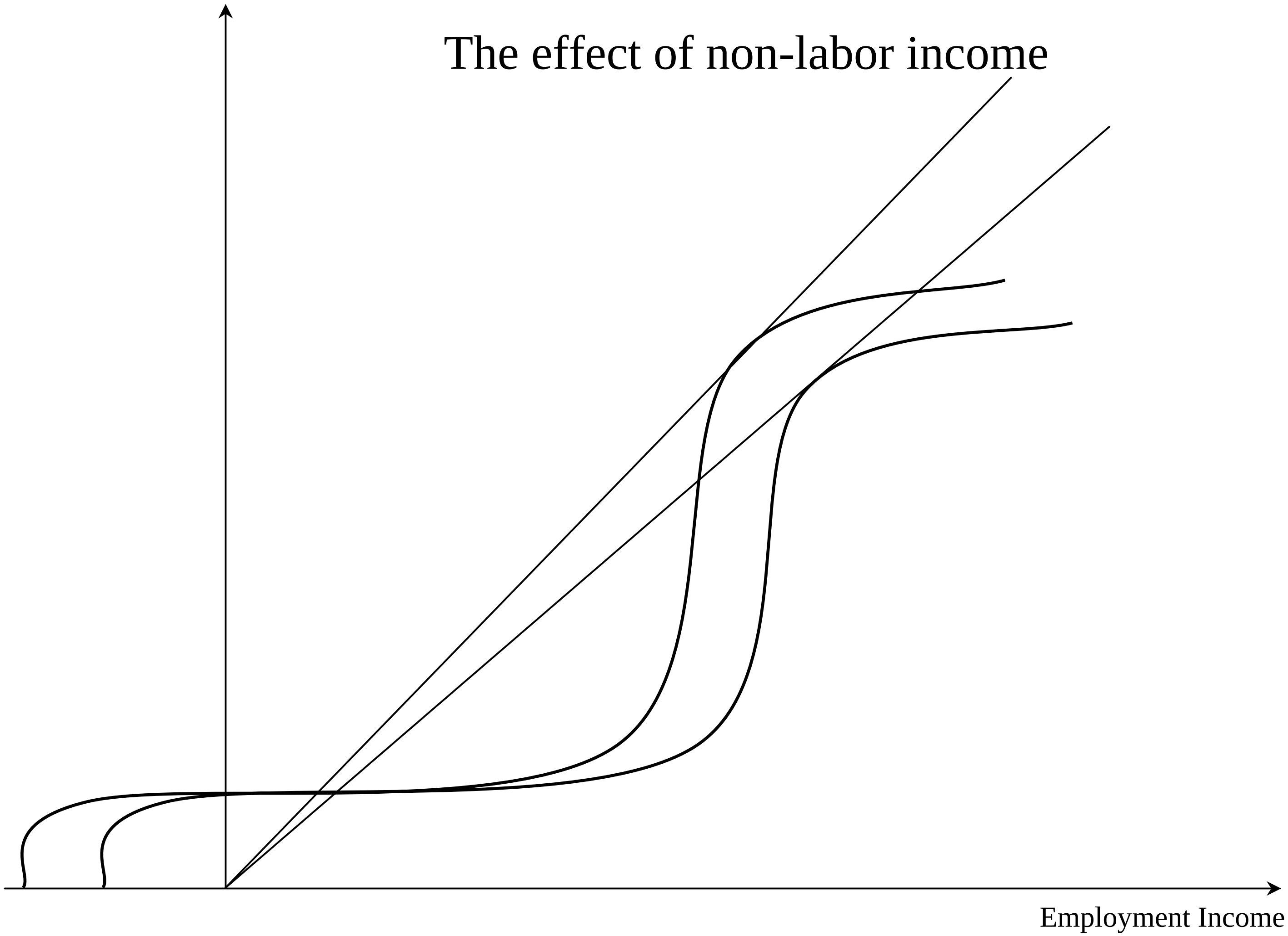


Figure 4: Effect of Non-Labor Income on the Capacity Curve

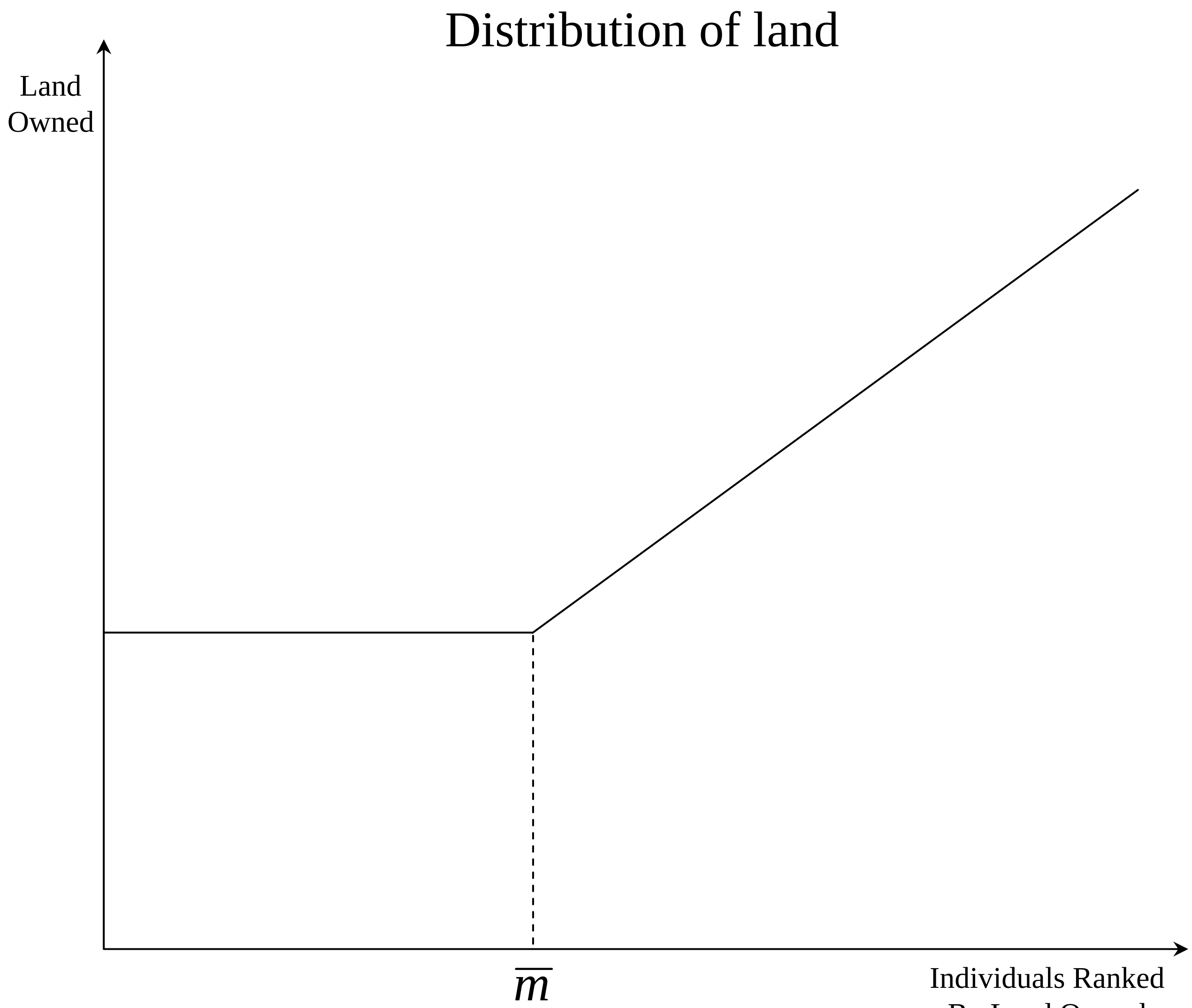


Figure 5: Distribution of Land

Labor supply as function of land owned

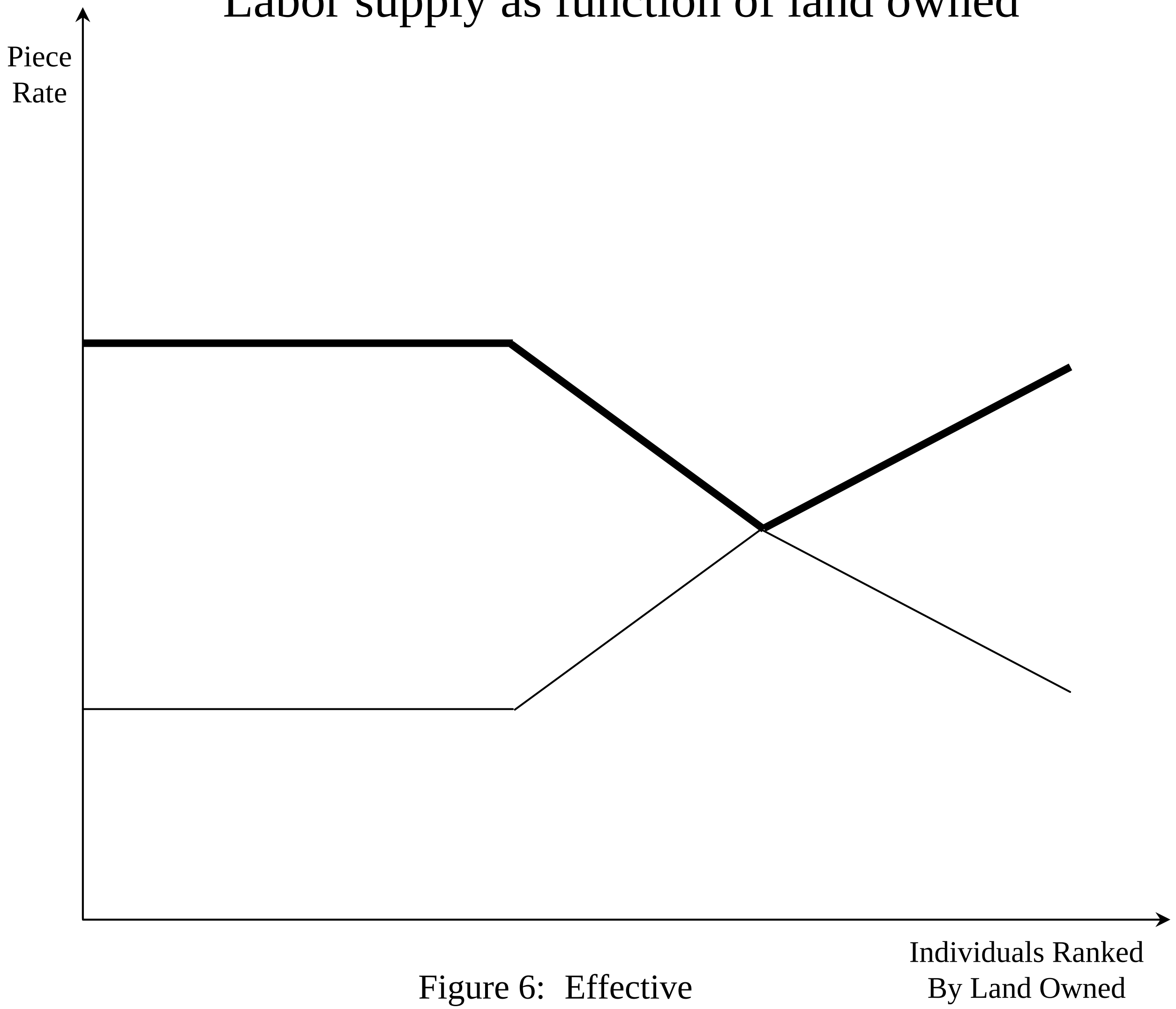


Figure 6: Effective Reservation Wage

Different types of equibria

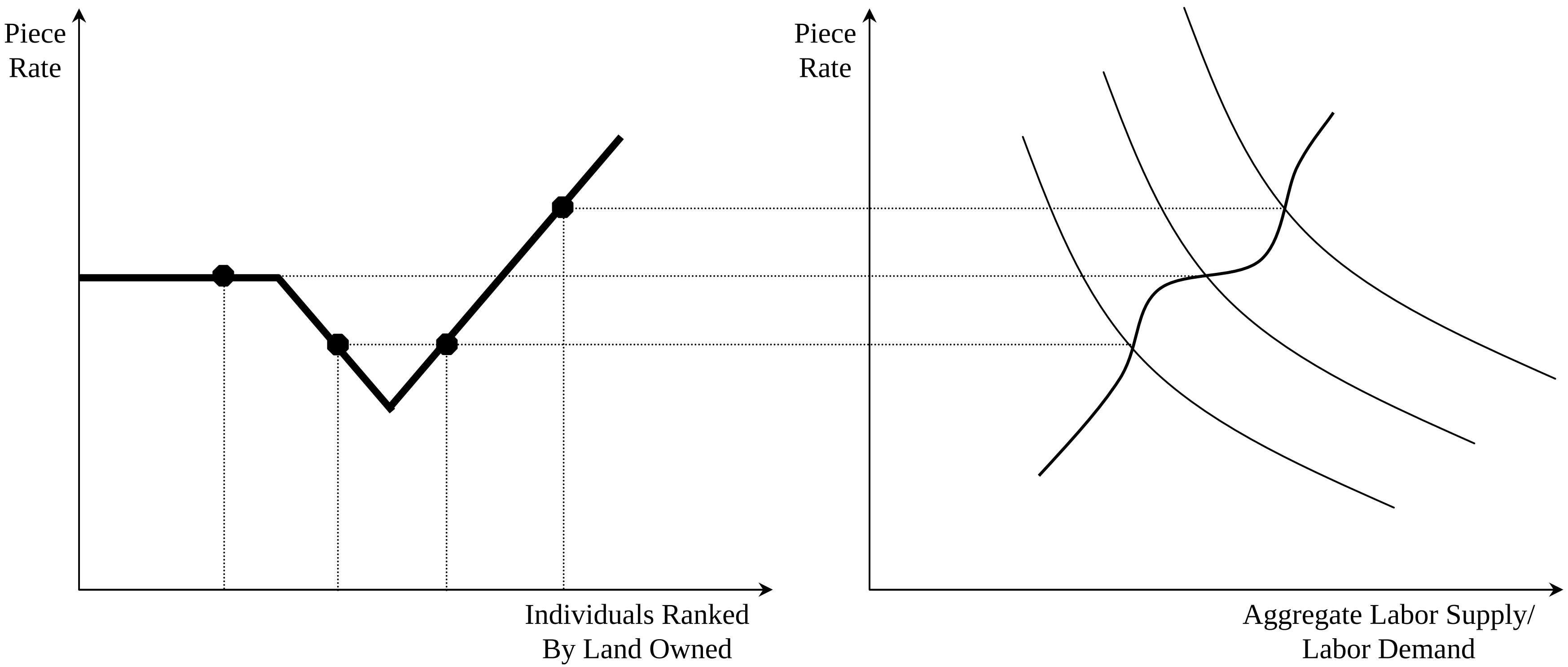


Figure 7: Type of Equilibria

Policy experiments

Dynamic versions of capacity curve

Dynamic version of the capacity curve

- Capacity curve:
  - Health affects income $$y_{t+1} = g(h_t)$$
  - Income affects health $$g(h_t) = f(y_t)$$

$$y_{t+1} = g(f(y_t))$$

- Capacity curve:

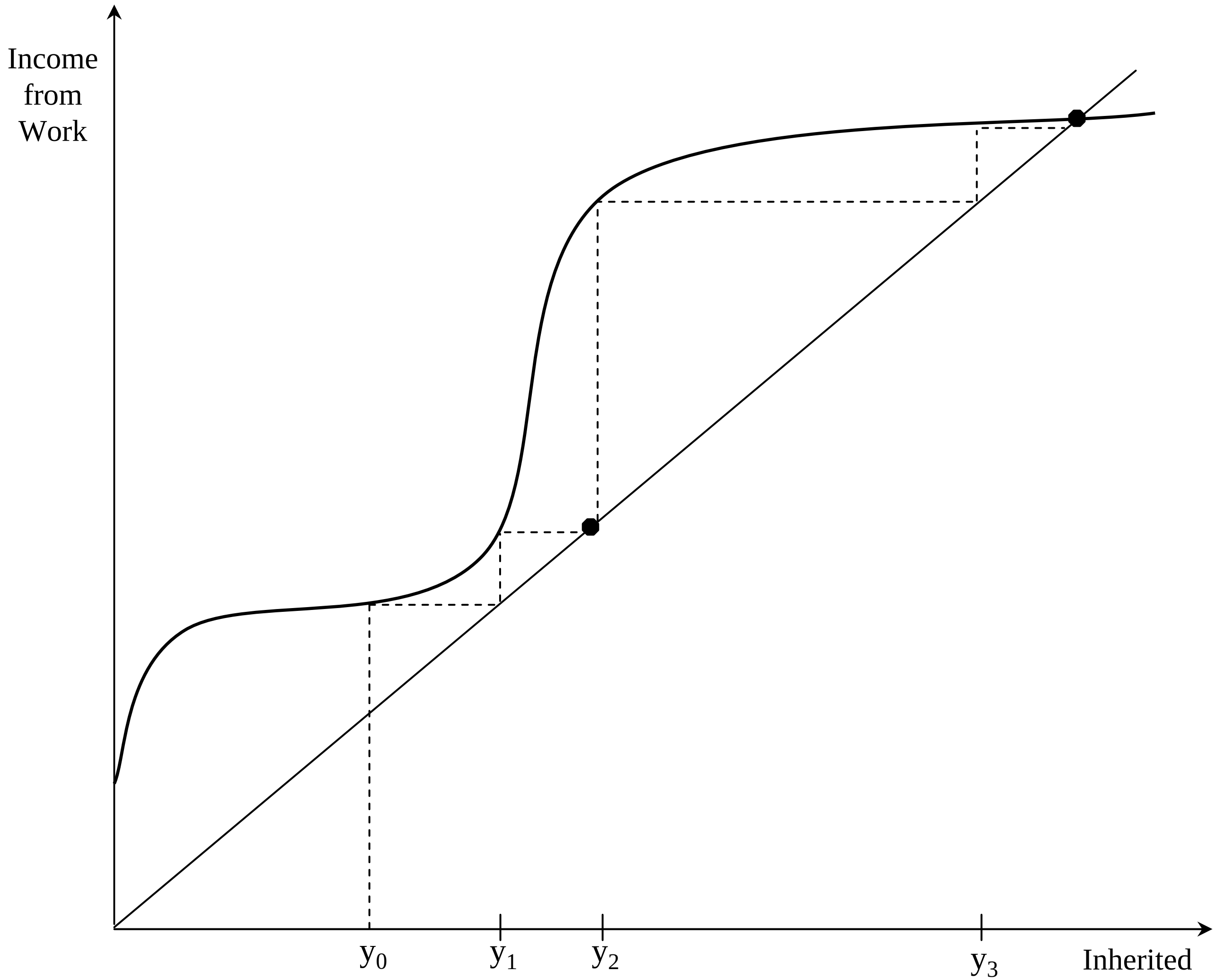


Figure 1

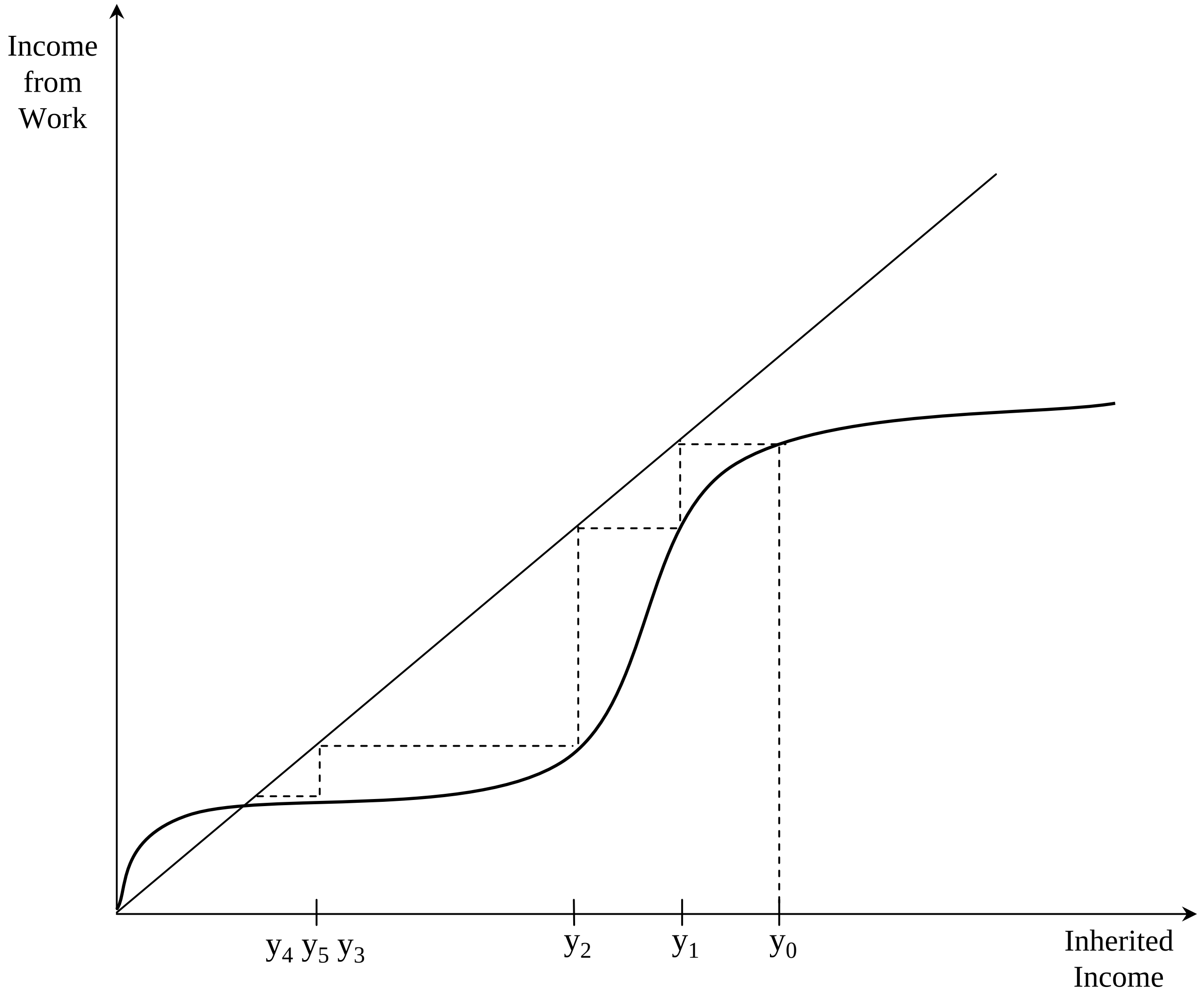


Figure 2

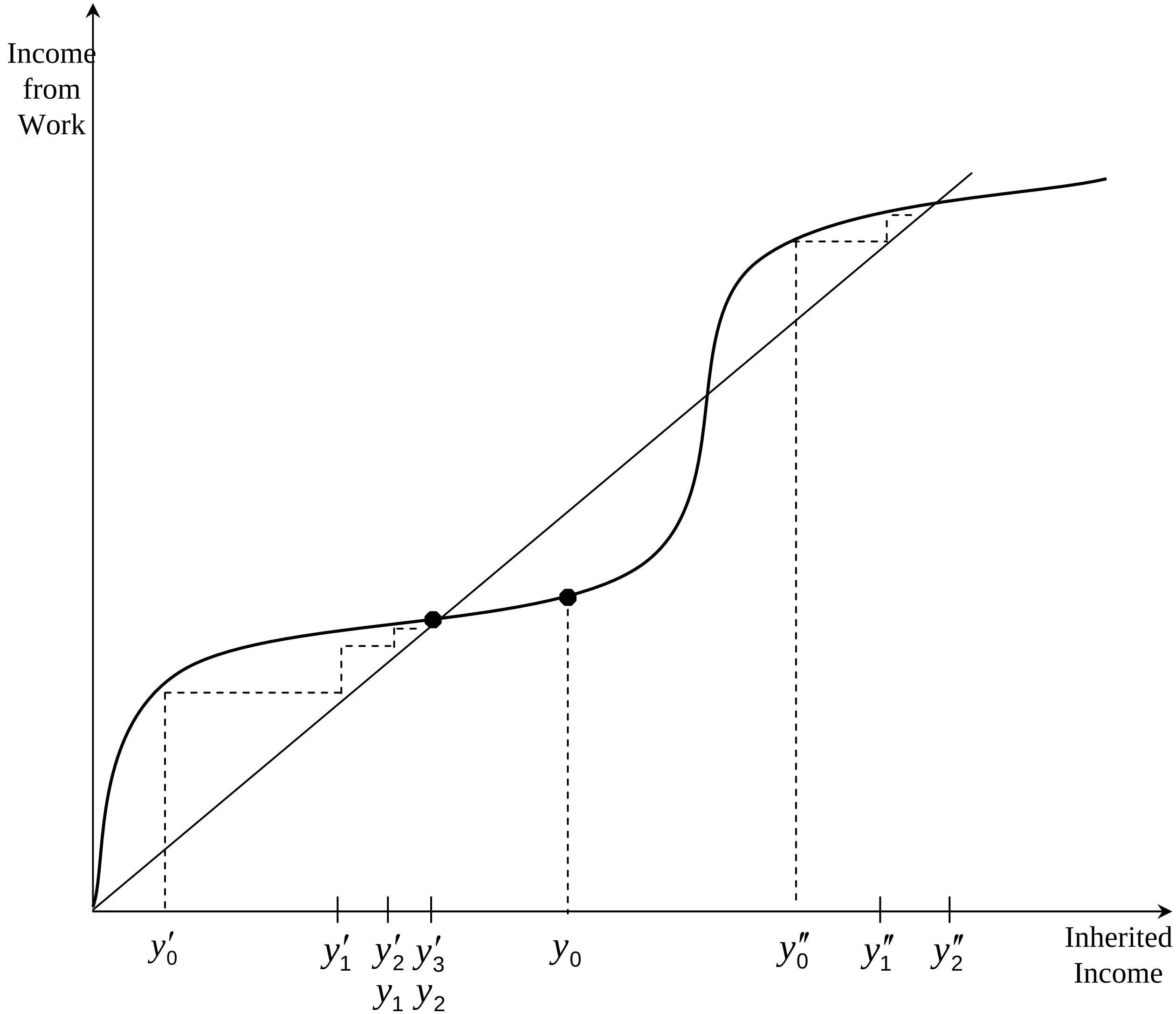


Figure 3

When will a property trap emerge?

- Capacity curve:
  - Health affects income $$y_{t+1} = g(h_t)$$
  - Income affects health $$g(h_t) = f(y_t)$$
- Capacity curve: $$y_{t+1} = g(f(y_t))$$
- Multiple equilibria (usually interpreted as poverty trap) will arise iff the capacity curve intersects the 45 degree line from below

Conditions for capacity curve to intersect 45 degree line from below

- Let *y** be the point at which the capacity curve intersect the 45 degrees line. At this point, the derivative $((g(f(y*))' > 1$

- Now: $(g(f(y*))' = g'(f(y*)) * f'(y*)$

$$= \frac{g'(f(y*))f(y*)}{g(f(y*))} \frac{f'(y*)y*}{f(y*)}$$

- Because $g(f(y*)) = y*$

- $\dfrac{g'(f(y*))f(y*)}{g(f(y*))}$ is the elasticity of *g* with respect to *h* (income with respect to health)
- $\dfrac{f'(y*)y*}{f(y*)}$ is the elasticity of *f* with respect to *y* (health with respect to income)
- By continuity, over some range, the product of the elasticities must be greater than one.
- A very general point, which we will now explore in the case of health

Caveats

- What is a period? (one day? One life time? One year?)
- What is health?
- What is income?

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

The Demand for Health and Calories

Esther Duflo

14.771

Some striking facts

- Achieving a minimum nutrition standard is fairly cheap (even with today's prices)
- Yet, there is plenty of malnutrition, even outside countries that face particularly acute food crises: ▸ Udaipur, ▸ Children malnutrition; ▸ Comparison with world ▸ calories consumed in India today.
- Household spend considerable share of their budget on health care, and visit doctors frequently: ▸ Visits to Doctors in Udaipur; ▸ health in the world ▸ Share of Budget spent on Health
- Yet, take up of cheap, highly effective preventive care remain really low:
  - Less than 5% of children and 3% of pregnant mothers in Kenya sleep under a bednet (Cohen and Dupas (2007)).
  - 25% of mother breastfeed within an hour of birth in India
  - NFHS report 44% of children fully immunized, and this is probably an exageration.

Understanding Health Demand

- Traditional model of demand for health input (e.g. Strauss and Thomas) (calories, an action, a shot)

$$H = f(y, p, r, X)$$

- For H, an health input. Health depends on price p , income y, characteristics and anticipated impact of these inputs on health r, X characteristics which may influence the demand for health itself (social norms, education, etc.)
- Underlying this model, view of health (and possibly health inputs as well) as both consumption and investment. Two sources for the income effect (health –or even health inputs–is a normal good; resource constraints).
- Health is difficult to understand: Education, environment, also play a role in determining r (perceived impact of action on health), and hence, in reduced form, demand for health inputs.
- This ignores intra-family decisions: whose health? who decides about it? who cares about it, to which we will return later.

Does Income affect Nutrition: The Engel-Curve approach?

- Deaton-Subramanian: Non-parametric estimates of the relationship between total expenditures per capita in the household and calories consumed.
- Points to keep in mind:
  - Food expenditure is not equal to nutrition: price per calory.
  - Need to go from food items consumed to nutrients consumed through conversion tables (need detailed items).
  - Some other adjustments needed: meals taken out or given away, waste.
  - This is not a structural relationship: it could be an income effect on the demand for food OR that the household needs more food to produce more income.
  - It is also cross-sectional: it may not be a good indication of how a person would react if their income increased (unobserved heterogeneity; adaptation to nutrition level)

- Methods to estimate $y = g(x)$
  - Kernel regression: for each point x along a grid,weighted mean of y in a neighborood of x (weights are Kernel weights)
  - Local Linear (Fan) regression: for each point x along a grid, predicted value from a weighted linear regression in a neighborhood of x (kernel weights). Better (not biased at the edges) (what Deaton and Subramanian use).
  - Further advantage: with y=log(calories per capita) and x=log(expenditure per capita), elasticity at any point is slope of the curve: directly estimated at each point in the grid

Results

- Clear relationship between total expenditures per capita and calorie consumption: ▸ figure
- The relationship does not appear to be non-linear, at least in this range, and the elasticity is never above 1 (despite the fact that it is probably an over estimate due to the reverse causality): ▸ Elasticity
- There is also a strong relationship between price of calories and expenditures (see ▸ figure, indicating a lot of substitution towards more expensive calories: not clear that households' back is against the wall, even very poor households.
- Since the relationship is more or less log-linear, they proceed to estimate a log-linear relationship, which allows them to add control variables: ▸ Table.

Deaton-Dreze Nutrition Over Time in India

- Surprisingly, calory consumption is falling down in India, at the same time as income and overall expenditure increase...
- Has the Engel-Curve gone away?
- No... instead it is shifting downwards over time: figure.
- How can we reconcile an upward-slopping Cross-Sectional Engel-Curve with a downward drifts as income increase?
  - Relative price of food?
  - Demand for variety
  - D-D explanation: Calory requirements are going down over time (better health infrastructure/more sedentary work); people do need to get enough calories to do their work, and try to get the tastier one possible within this constraint
  - Problem with that explanation: nutritional status is still really poor in India, and not going up really fast (in NFHS some measures show a worsening of children's nutritional status). This explanation assumes that people are content with a nutritional status that most of us would consider quite poor.

Income shock and nutritional status: Natural Experiments

- Indian time-series experience suggest that movements along the cross-sectional Engel curve may not identify the effect of income on demand for nutrition
- Time series does not identify this impact either, since many things change over time.
- Ideal (thought) experiment?
- Natural experiments:
  - Duflo: Grandmother and Granddaughters: exogenous pension increase in South Africa on child health (we will discuss it later when we talk about intra-household).
  - Banerjee, Duflo, Postel-Vinay, Watts: 40% drop in wine production (due to phyloxera). Income shock with probably not too many other things happening. Look at heights of cohorts born at that time: table.

The demand for Calories: Price effect

- Large increase in food prices since 2005. From March 2007 to March 2008, the average world price for corn increased 30%; for rice, it increased 74%; for soybeans, 87%; and for wheat, 130%
- Very concerning for the welfare of the poor.
- Will it result in stark decline in nutritional standards?
- Note that an increase in prices will have a both an income and a substitution effect (since food is an important part of the budget).
- Income effect should lead to a substitution towards cheaper food items (even if they all increase proportionally)

The demand for Calories: Jensen-Miller

- Reverse Experiment in China: subsidize staple food in two region for randomly selected household. Survey food consumption after a few month.
- In both regions, substitution towards more expensive calories: ▸ Hunan ▸ Guansu.
- In one region, calories consumption actually worsens. No perceptible improvement on the other items except fat. In the other region, no change in calories consumption ▸ Table.
- What can explain these results?
- Caveats: short term decrease in food prices: people may be using the windfall to have good food rather than to improve their nutritional status. Long term increase/decrease may have very different impacts.

The demand for Health: Price effects

- Households seem to be willing to pay for what they perceive to be quality: e.g. expenditure on private health care in India.
- However, surprisingly large price elasticities for preventive health–Examples:
- Small positive price: Cohen-Dupas: Take up of insecticide treated bednets by pregnant women
  - Experimental Approach: Different maternity clinics randomly assign to give away nets or to sell them at different prices.
  - Huge elasticity of take up: figure
  - No elasticity of use conditional on take up (contrary to what is often hypothesized): Conditional usage Effective coverage.
  -

Price Effects

- Small negative price: Banerjee-Duflo etc.. Immunization in India
  - 130 villages
  - 60 get randomly assigned to receive regular immunization camp
  - 30 of those get small incentives to get immunized (1 kg of lentils/1 set of plates)
  - Pretty large impact of the camp..But larger impact of the lentils ▸ complete immunization ▸ number of shots

Why these high price elasticities?

- These are just two examples, but Kremer and Holla (2008) review several
- High price elasticities, particularly for preventive health.
- Basis for the "Conditional Cash transfer approach" (e.g. Progresa, Mexico), which has become very popular in many countries: cash transfer is conditional on health.
- Puzzling in light of health demand model we started with:
  - Large benefits
  - Prices (or opportunity cost) are not that high to begin with
- High discount rates or hyperbolic discounting?
- Problem with this explanation is that it would imply a pretty high degree of naivete to keep postponing (given the large benefits)

High elasticity: possible Explanation

- We observe large sensitivity to relevant information
  - e.g. Dupas (2007): Pregnancies with older partners decline significantly when teenager are informed that older men are less likely to have HIV than younger men. ▸ table
- However, learning about health is very difficult (Das and Sanchez, 2002): many diseases are "self-limiting", in the sense that symptoms will go away by themselves (at least temporarily) : learning about doctor quality is really hard.
- Particularly difficult to link cause and effects with preventive case, especially when the behavior has externalities.

High elasticity: possible Explanation

- Mistrust of government (message changes; in India: forced sterilization).
- Very little impact of less well defined "behavioral change" message
  - e.g. Kremer and Miguel: No impact of a campaign to convince kids to wear shoes and stop fishing in the lake to avoid catching worms.
- As a result, there could be a large zone of indifference, where people simply do not know what the benefits are, so are willing to do things if this does not cost them anything, but will be discouraged by any cost.
- May not be the right model... However, what this does suggest is that the basic "health as human capital" model is problematic.

Conclusions

- At a literal level, the Das Gupta - Ray model is not doing very well so far... the income elasticities of calory consumption are probably not zero, but clearly not huge either.
- However, this may need to be re-interpreted less literally.
- May be health rather than income.
- Think of the very high price elasticity for the very poor. If those were much lower for the rich, small differences in the political environment could generate huge difference in health outcomes for rich and poor.

Banerjee, Deaton, and Duflo (Health, Wealth, Health Services), Table 1

Selected Health Indicators, by Position in the per Capita Monthly Expenditure Distribution

Indicator	*Group*		
	Bottom third	*Middle third*	*Top third*
Reported health status	5.87	5.98	6.03
No. symptoms self-reported in last 30 days	3.89	3.73	3.96
BMI	17.85	17.83	18.31
Hemoglobin below 12 g/dl	0.57	0.59	0.51
Peak flow meter reading	314.76	317.67	316.39
High blood pressure	0.17	0.15	0.20
Low blood pressure	0.06	0.08	0.09

Notes: Means reported are based on data collected by the authors from 1,024 households. See text for survey and variable description.

Figure by MIT OpenCourseWare.

Deaton and Dreze, Table 9

Child Nutrition Indicators, 1975-9 to 2004-5

	Proportion (%) of undernourished children 1975-79	1988-90	1996-97	2000-1	2004-5	Percentage decline (1975-9 to 2004-5)
Weight-for-age						
Below 2 SD	77	69	62	60	55	29
Below 3 SD	37	27	23	21	18	51
Height-for-age						
Below 2 SD	79	65	58	49	52	34
Below 3 SD	53	37	29	26	25	53
Weight-for-height						
Below 2 SD	18	20	19	23	15	17
Below 3 SD	2.9	2.4	2.5	3.1	2.4	17
Prevalence of nutritional deficiency signs (%)						
Oedema						
Marasmus	0.4	0.1	0.1	0.0	0.0	100
Bitot spots	1.3	0.6	0.1	0.2	0.0	100
Angular stomatitis	1.8	0.7	0.7	0.8	0.6	67
	5.7	5.7	2.1	1.4	0.8	86

Source: *National Nutrition Monitoring Bureau (1991, 1999, 2002, 2006). All figures pertain to children aged 1-5 years.*

Figure by MIT OpenCourseWare.

Deaton and Dreze, Table 10

Countries with the Highest Levels of Child Undernutrition, 1996-2005

Country	Proportion (%) of children with low "weight for age"
Nepal	48.3
Bangladesh	47.5
India	46.7
Timor-Leste	45.8
Yemen	45.6
Burundi	45.1
Madagascar	41.9
Sudan	40.7
Lao (People's Dem Rep)	40.4
Niger	40.1
Eritrea	39.6
Afghanistan	39.3

Source: *World Development Indicators, 2007. Figures apply to the most recent year for which data are available within the reference period. There is a significant margin of error for individual countries.*

Figure by MIT OpenCourseWare.

Deaton and Dreze, Table 5

Fractions of the Population Living in Households with per Capita Calorie Consumption below 2,100 Urban and 2,400 Rural

Year	Round	Rural	Urban	All India
1983	38	66.1	60.5	64.8
1987-8	43	65.9	57.1	63.9
1993-4	50	71.1	58.1	67.8
1999-0	55	74.2	58.2	70.1
2004-5	61	79.8	63.9	75.8

Source: *Authors' calculations based on NSS data.*

Figure by MIT OpenCourseWare.

Banerjee, Deaton, and Duflo (Health Care Delivery), Table 2

Frequency of Health Care Visits

Per capita monthly expenditure		*Total number of visits in the last 30 days*			
		All	*Public*	*Private*	*Bhopa*
Panel A: Means					
All	470	0.51	0.12	0.28	0.11
Poor	219	0.43	0.09	0.22	0.12
Middle	361	0.54	0.11	0.29	0.13
Rich	770	0.55	0.15	0.33	0.07
Panel B: OLS regressions: dependent variable: number of visits					
Middle		0.11	0.02	0.07	0.01
		(.052)	(.023)	(.034)	(.027)
Rich		0.12	0.06	0.11	-0.05
		(.05)	(.024)	(.034)	(.022)
Panel C: OLS regressions, with village fixed effects					
Middle		0.14	0.02	0.09	0.02
		(.047)	(.024)	(.033)	(.023)
Rich		0.13	0.04	0.11	-0.03
		(.05)	(.026)	(.036)	(.025)
Villages fixed effects		yes	yes	yes	yes

Note: Omitted dummies in panel B and C: poor
Standard errors in parentheses below the coefficients

Figure by MIT OpenCourseWare.

Health in the Household

	In Last Month				
	Percent of HH Members Sick	A Household's Average # of Consultations	Percent of Households that met At Least Once with a Consultant: Public	Private	Infant Mortality
Rural					
Cote d'Ivoire	21.4%	1.28	49.7%	3.2%	6.2%
Guatemala					6.2%
India - Hyderabad					
India - Udaipur	46.1%	0.11	20.1%	58.1%	10.0%
India - UP/Bihar	12.5%	0.81	13.9%	47.3%	7.7%
Indonesia	24.2%	0.77	20.7%	27.3%	3.4%
Mexico	46.3%	1.11	47.7%	0.0%	6.9%
Nicaragua	34.9%	0.15	46.0%	5.0%	
Pakistan	28.0%	0.45	24.0%	48.8%	16.7%
Panama	15.2%	0.10	23.8%	0.0%	
Papua New Guinea					
Peru	11.1%	0.10	20.9%	8.5%	
South Africa	12.5%	0.12	16.4%	6.9%	8.6%
Tanzania	13.2%	0.07	23.2%	14.0%	8.7%
Timor Leste	11.7%	0.21	30.2%	0.5%	

How the poor spend their money

As a Share of Total Consumption

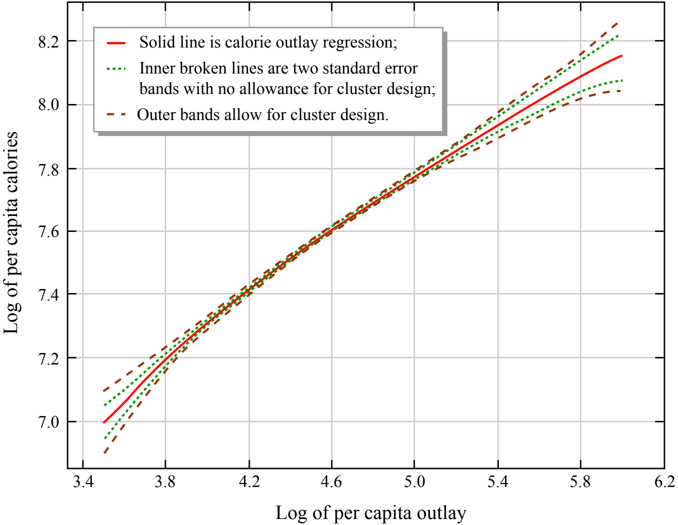
	Food	Alcohol/ Tobacco	Education	Health
Rural				
Cote d'Ivoire	64.4%	2.7%	5.8%	2.2%
Guatemala	65.9%	0.4%	0.1%	0.3%
India - Udaipur	56.0%	5.0%	1.6%	5.1%
India - UP/Bihar	80.1%	3.1%	0.3%	5.2%
Indonesia	66.1%	6.0%	6.3%	1.3%
Mexico	49.6%	8.1%	6.9%	0.0%
Nicaragua	57.3%	0.1%	2.3%	4.1%
Pakistan	67.3%	3.1%	3.4%	3.4%
Panama	67.8%		2.5%	4.0%
Papua New Guinea	78.2%	4.1%	1.8%	0.3%
Peru	71.8%	1.0%	1.9%	0.4%
South Africa	71.5%	2.5%	0.8%	0.0%
Timor Leste	76.5%	0.0%	0.8%	0.9%

Deaton and Subramanian, Table 1

Expenditure Patterns, Calorie Consumption, and Prices per Calorie, Rural Maharashtra, 1983

	Expenditure shares (%)			*Calorie shares (%)*			*Price per calorie (Rupees per 1,000 calories)*		
	Mean (1)	*Bottom 10% (2)*	*Top 10% (3)*	*Mean (4)*	*Bottom 10% (5)*	*Top 10% (6)*	*Mean (7)*	*Bottom 10% (8)*	*Top 10% (9)*
A. Food groups									
Cereals	40.7	46.0	31.0	70.8	77.3	57.3	.64	.51	.79
Pulses	8.9	10.2	7.8	6.6	6.2	7.2	1.51	1.44	1.60
Dairy	8.1	4.9	11.8	2.8	1.3	4.9	3.69	3.59	3.92
Oils and fats	9.0	9.2	9.2	5.9	4.8	7.6	1.74	1.67	1.81
Meat	5.1	3.4	6.4	.7	.4	1.0	11.7	11.0	12.2
Fruits and vegetables	10.5	8.5	12.0	3.5	2.3	5.4	3.90	3.83	3.85
Sugar	6.5	7.4	5.9	7.2	7.0	8.0	1.01	.94	1.09
Other food	11.3	10.4	16.1	2.5	0.8	8.6	17.4	16.8	15.9
B. Cereals									
Rice	11.6	9.0	10.9	15.2	10.1	16.5	.95	.89	1.02
Wheat	5.6	3.8	7.9	8.5	4.7	14.4	.79	.73	.82
Jowar	18.2	27.4	9.3	37.8	52.9	21.6	.50	.43	.55
Bajra	3.0	2.7	1.3	6.6	4.9	3.2	.48	.48	.50
Other coarse cereal	1.2	2.8	.3	2.2	4.5	.6	.66	.58	.99
Cereal substitutes	1.1	.5	1.3	.6	.2	.8	2.23	2.22	2.22
Total food (or total calories)	67.4	73.4	54.1	2,120	1,385	3,382	1.14	.88	1.50
				2,098	1,429	3,167			

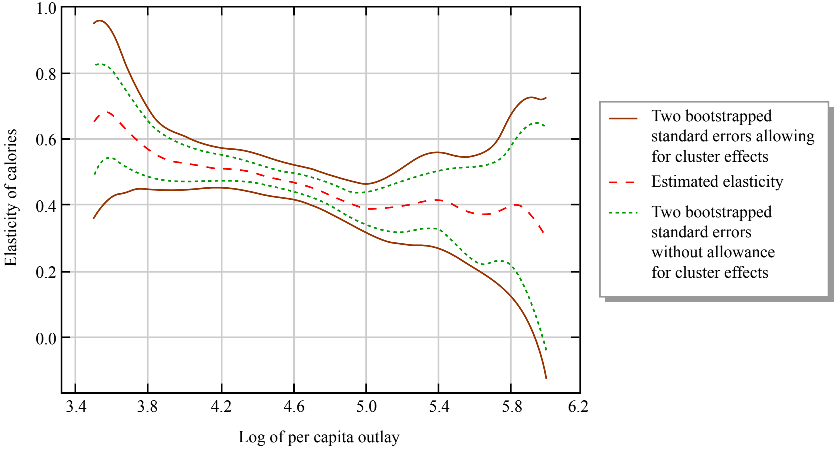
Figure by MIT OpenCourseWare.



Regression function for log calories and log per capita expenditure, Maharashtra, India, 1983.

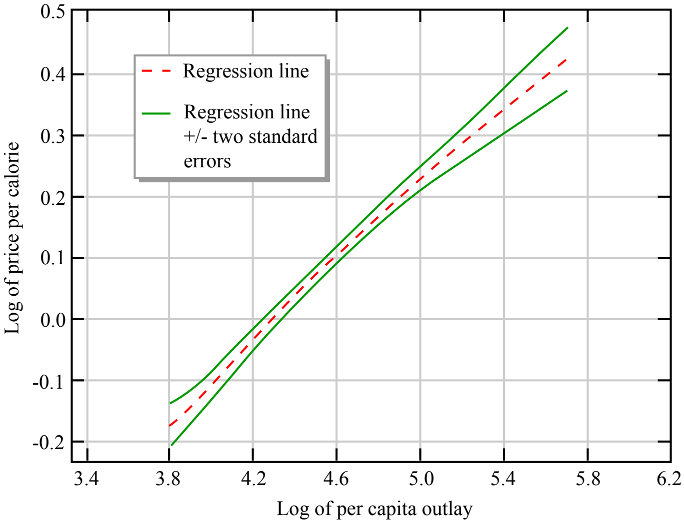
Figure by MIT OpenCourseWare.

Deaton and Subramanian, Figure 3



Elasticity of per capita calories to per capita expenditure, Maharashtra, India, 1983.

Figure by MIT OpenCourseWare.



Log of price per calorie and log of per capita expenditure, Maharashtra, India, 1983.

Figure by MIT OpenCourseWare.

OLS Estimates of Double Log Calorie and Calorie Price Regressions with Other Covariates

	Log calorie availability				Log price per calorie			
	All data (1)		Within village (2)		All data (3)		Within village (4)	
	β	$\|t\|$	β	$\|t\|$	β	$\|t\|$	β	$\|t\|$
Constant	6.028	(78)			- 1.5934	(18)		
ln PCE	.3655	(29)	.3407	(27)	.3799	(25)	.3217	(23)
ln household size	- .1572	(14)	- .1630	(21)	.0839	(6.8)	.0661	(8.4)
rm04	- .0967	(2.2)	- .1461	(4.1)	.1024	(2.3)	.1008	(3.3)
rm59	.0488	(1.2)	.0321	(1.0)	- .0467	(1.2)	- .0331	(1.2)
rm1014	.0891	(1.9)	.0612	(1.9)	- .1120	(2.3)	- .0842	(2.9)
rm1555	.1636	(5.1)	.1634	(5.9)	- .1700	(4.3)	- .1347	(5.0)
rm55+	.1406	(3.0)	.1213	(2.8)	- .1565	(3.6)	- .1074	(2.9)
rf 04	- .1359	(3.1)	- .1869	(4.9)	.0460	(1.1)	.0742	(2.2)
rf 59	.0176	(.4)	- .0040	(.1)	- .0643	(1.4)	- .0476	(1.4)
rf 1014	.1140	(2.8)	.0679	(2.0)	- .1108	(2.7)	- .0873	(3.0)
rf 1555	.0420	(1.6)	.0514	(2.1)	.0085	(.3)	- .0021	(.1)
Scheduled caste	- .0083	(.8)	- .0179	(2.0)	.0020	(.2)	- .0071	(.8)
Hindu	.0114	(.7)	.0302	(2.1)	- .0562	(2.6)	- .0605	(4.4)
Buddhist	.0237	(1.1)	.0400	(2.0)	- .1080	(4.0)	- .0760	(4.0)
Self-employed nonagriculture	.0187	(1.0)	.0064	(.4)	- .0270	(1.1)	.0079	(.5)
Agricultural labor	.0433	(2.2)	.0222	(1.4)	- .0837	(3.4)	- .0418	(2.7)
Nonagricultural labor	.0275	(1.1)	.0293	(1.5)	- .0210	(.8)	- .0315	(1.7)
Self-employed agriculture	.0618	(3.5)	.0389	(2.7)	- .0610	(2.8)	- .0118	(.8)
R^2	.5532		.6706		.4254		.6414	

Figure by MIT OpenCourseWare.

Deaton and Dreze, Figure 1

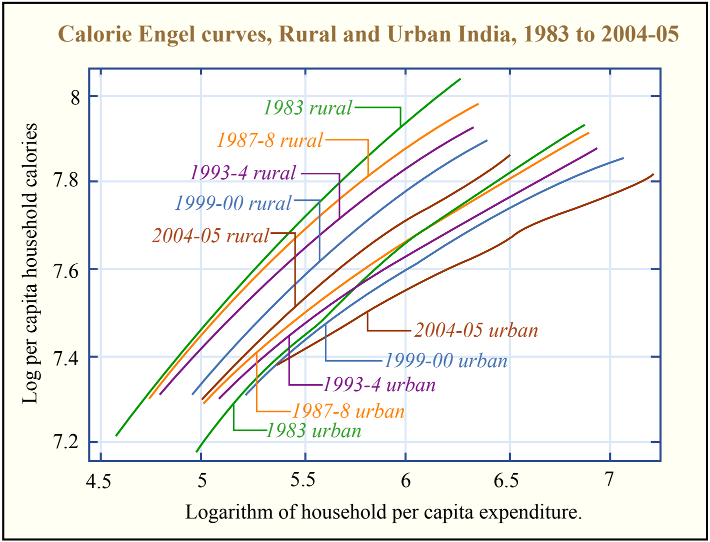


Figure by MIT OpenCourseWare.

Table 3: Impact of phylloxera on height at 20
(year of birth 1852-1892: cohorts 1872-1912)

	Dependent Variables	
	Mean height	Fraction shorted than 1.56 meter
	(1)	(2)
A. Year Dummies, Departement dummies		
Born in phylloxera year	-0.00150	0.00358
	(.00093)	(.00204)
Observations	3485	3485
Department trend	No	No
B.Year dummies, departement dummies, departement trend		
Born in phylloxera year	-0.00188	0.00381
	(.00095)	(.00173)
Observations	3485	3485
Department trend	Yes	Yes
C. Hectare of vine per habitant		
born in phylloxa year *hectare	-0.00753	0.01928
vine per habitant	(.00389)	(.01142)
Observations	3485	3485
Department trend	Yes	Yes

Table 4. Consumption Response to the Price Subsidy

	HUNAN							
	Rice	Other Cereal	Fruit & Veg	Meat	Seafood	Pulses	Dairy	Fats
%Subsidy(rice)	-0.235*	0.397	-0.623***	0.377	0.482**	-0.791*	-0.054	-0.567*
	(0.140)	(0.355)	(0.227)	(0.415)	(0.230)	(0.476)	(0.069)	(0.313)
%ΔEarned	0.043***	-0.001	0.058***	0.002	0.036	-0.052	-0.006	0.022
	(0.014)	(0.040)	(0.021)	(0.043)	(0.022)	(0.050)	(0.004)	(0.031)
%ΔUnearned	-0.044*	-0.087	-0.018	0.076	-0.004	-0.037	-0.021	-0.007
	(0.025)	(0.065)	(0.040)	(0.071)	(0.042)	(0.075)	(0.019)	(0.055)
%ΔPeople	0.89***	0.46**	0.63***	0.05	-0.07	0.48**	0.09	0.88***
	(0.08)	(0.19)	(0.11)	(0.24)	(0.10)	(0.23)	(0.05)	(0.16)
Constant	4.1***	7.5***	-0.3	-5.7**	-0.2	8.8***	0.2	-8.3***
	(1.0)	(2.5)	(1.4)	(2.8)	(1.4)	(3.0)	(0.6)	(2.1)
Observations	1258	1258	1258	1258	1258	1258	1258	1258
R^2	0.19	0.06	0.11	0.07	0.02	0.03	0.02	0.09

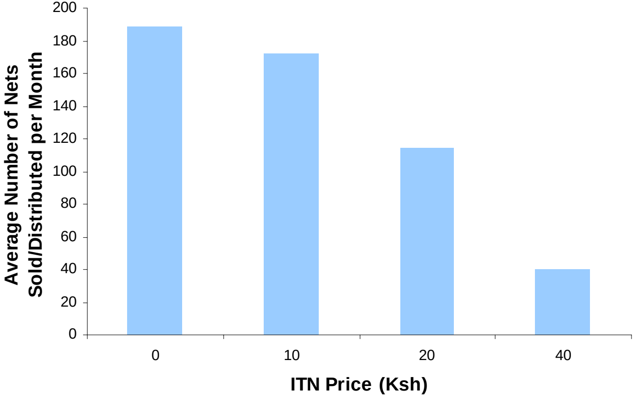
					GANSU			
	Wheat	Other Cereal	Fruit & Veg	Meat	Seafood	Pulses	Dairy	Fats
%Subsidy(wheat)	0.353	-0.283	0.049	0.130	-0.017	0.240	0.282	0.507**
	(0.258)	(0.335)	(0.190)	(0.299)	(0.017)	(0.320)	(0.207)	(0.251)
%ΔEarned	0.079**	-0.067	0.061**	0.085*	0.000	-0.047	-0.025	0.091***
	(0.036)	(0.049)	(0.027)	(0.044)	(0.000)	(0.043)	(0.029)	(0.033)
%ΔUnearned	-0.017	0.130	0.046	0.314***	0.025	0.012	0.108	-0.110
	(0.092)	(0.106)	(0.077)	(0.091)	(0.025)	(0.104)	(0.073)	(0.091)
%ΔPeople	0.58***	0.52*	1.01***	-0.10	-0.01	0.44**	0.10	0.66
	(0.22)	(0.29)	(0.15)	(0.28)	(0.01)	(0.18)	(0.12)	(0.15)
Constant	-26.1***	23.8***	11.0***	2.4	-0.2	6.0**	-3.4*	7.2
	(2.3)	(2.8)	(1.6)	(2.5)	(0.2)	(2.6)	(1.9)	(2.1)
Observations	1269	1269	1269	1269	1269	1269	1269	1269
R^2	0.08	0.06	0.07	0.05	0.03	0.06	0.03	0.07

Table 2. Calorie Response to the Price Subsidy

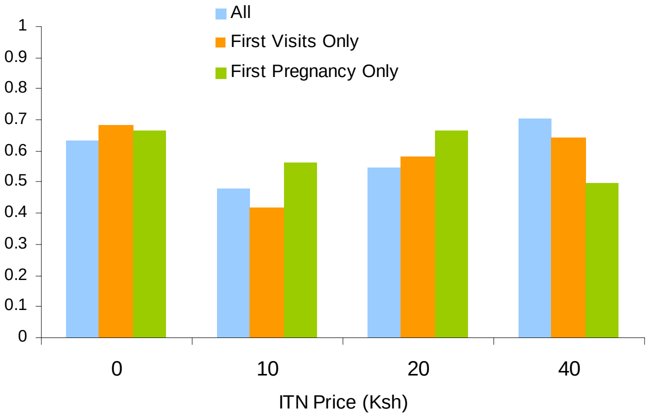
	HUNAN					GANSU				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Full Sample (Calories)	Below Median (Calories)	Above Median (Calories)	Bottom Quartile (Calories)	Full Sample (Protein)	Full Sample (Calories)	Below Median (Calories)	Above Median (Calories)	Bottom Quartile (Calories)	Full Sample (Protein)
%Subsidy(rice/wheat)	-0.206^{*}	-0.042	-0.339^{**}	0.004	-0.096	0.154	0.169	0.132	0.070	0.091
	(0.108)	(0.144)	(0.164)	(0.207)	(0.133)	(0.100)	(0.143)	(0.138)	(0.261)	(0.112)
%ΔEarned	0.031^{***}	0.026^{*}	0.036^{**}	0.037^{*}	0.040^{***}	0.028^{**}	0.027	0.029	0.053	0.017
	(0.011)	(0.014)	(0.018)	(0.021)	(0.013)	(0.014)	(0.021)	(0.019)	(0.034)	(0.016)
%ΔUnearned	-0.022	-0.025	-0.023	-0.037	-0.010	0.046	0.020	0.071^{*}	0.101	0.069
	(0.020)	(0.027)	(0.028)	(0.034)	(0.023)	(0.035)	(0.056)	(0.043)	(0.119)	(0.033)
%ΔPeople	0.94^{***}	1.07^{***}	0.80	1.04^{***}	0.93^{***}	0.91^{***}	1.01^{***}	0.81^{***}	1.08^{***}	0.88^{***}
	(0.07)	(0.08)	(0.11)	(0.10)	(0.07)	(0.08)	(0.10)	(0.12)	(0.13)	(0.09)
Constant	0.9	1.6	0.5^{***}	2.8^{*}	0.8	-1.9	0.1	-3.9	0.6	-4.0
	(0.8)	(1.1)	(1.1)	(1.5)	(0.9)	(0.8)	(1.1)	(1.1)	(1.7)	(0.9)
Observations	1258	633	625	317	1258	1269	634	635	320	1269
R^2	0.26	0.34	0.21	0.39	0.20	0.18	0.23	0.15	0.29	0.16

Notes: Regressions include county*time fixed-effects. The dependent variable in columns 1-4 and 6-9 is the arc percent change in household caloric intake and in columns 5 and 10 it is the arc percent change in household protein consumption. Standard errors clustered at the household level. %Subsidy (rice/wheat) is the rice or wheat price subsidy, measured as a percentage of the average price. %ΔEarned is the arc percent change in the household earnings from work; %ΔHH Unearned is the arc percent change in the household income from unearned sources (government payments, pensions, remittances, rent and interest from assets); %ΔPeople is the arc percent change in the number of people living in the household. *Significant at 10 percent level. **Significant at 5 percent level. ***Significant at 1 percent level.

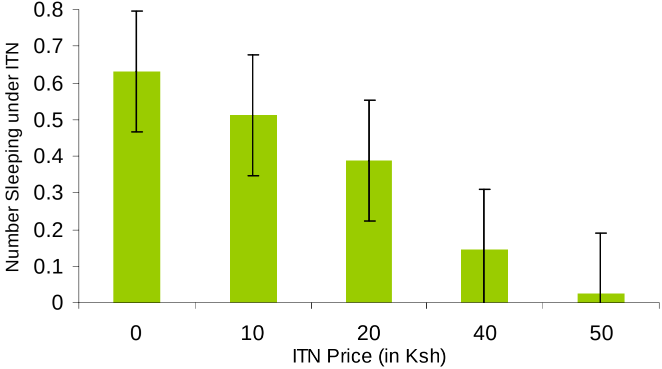
Results: Demand
Monthly Net Sales by ITN Price



Results: Usage Share Observed Using ITN at follow-up



Effective Coverage: Share of Prenatal Clients Sleeping Under ITN, by Price



Effective Coverage: Share of Prenatal Clients Sleeping Under ITN, by Price

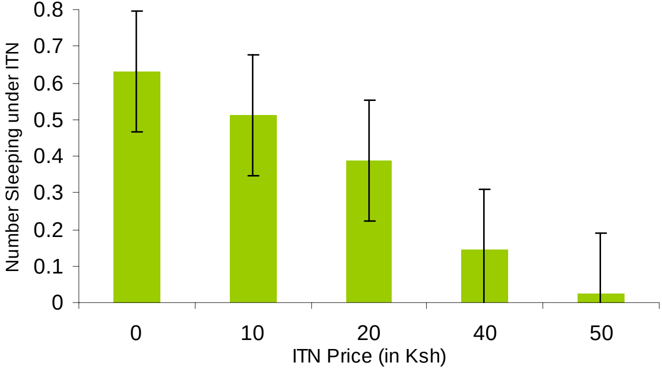
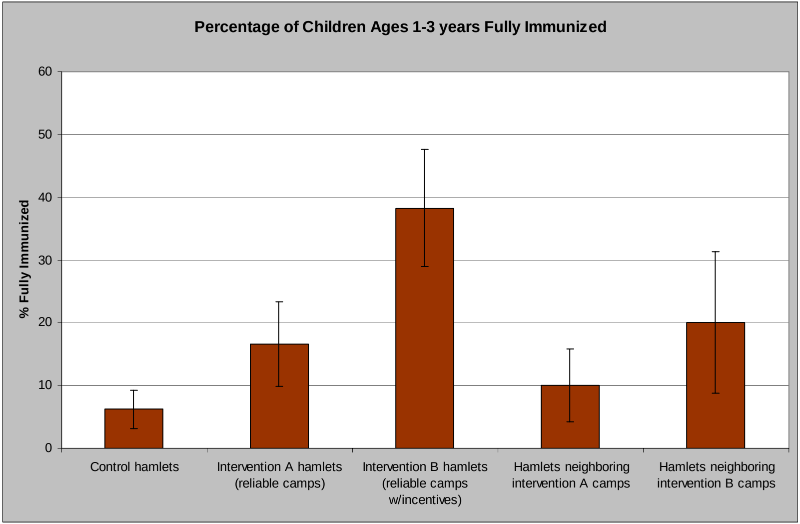


Figure 2: Percentage of children 1-3 years fully immunized by intervention status



Note: Fully immunized is defined as reporting 5 or more immunizations. Weighted means are reported, and the bars reflect the 95% clustered confidence interval.

Figure 3: Number of immunizations received by children 1-3 years

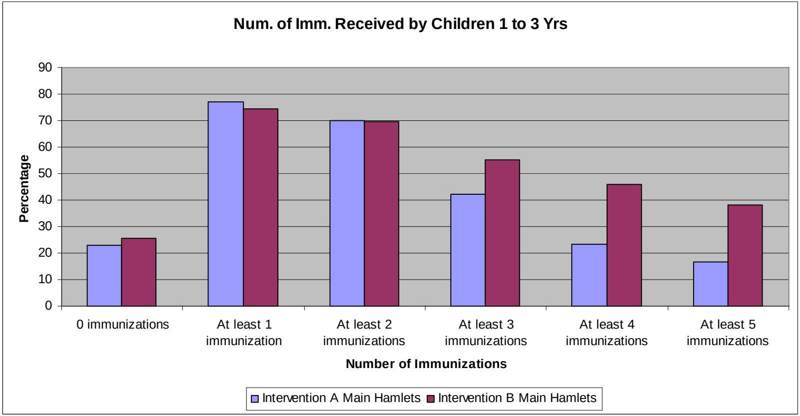


Table 7
Overall Treatment Effect on Incidence of Childbearing by Adult Men

	Comparison Group Base=100	Treatment Group	# Averted	Treatment Effect
# Observed Teen Pregnancies	100	68.6	31.4	-31.4%
Share of Observed Pregnancies by Adult Men	48%	24%		-23.2%
# Observed Pregnancies by Adult Men	47.6	16.7	30.9	-64.8%
# Observed Pregnancies by Young Men	52.4	51.9	0.5	-1.0%
Share of Cross-Generational Pregnancies among Averted Pregnancies			98%	

Notes: Treatment = Relative Risks Information Campaign. First row: treatment effect on number of teen pregnancies reported from Table 5 (-0.17/0.53). Second row: treatment effect on share of pregnancies by adult men reported from Table 6, regression (3).

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

Does Health Affect Productivity?

Esther Duflo

14.771

The Impact of Health on Productivity?

- Strong biological reasons to think that health (and nutrition) affects productivity: strength, days of illness, etc.
- At the micro-level, some indicators of health show fairly strong relationship with earnings ▸ StraussFigure
- At the macro-level, some have argued extremely high impact of health on GDP per capita (Sachs, Commission on macroeconomics and health)
- E.g. Gallup and Sachs (2003) log(GDP/capita)=-1.3*Population exposed to Malaria
- Potential problems with these estimates (both micro and macro?)
- Today we will focus on both micro and macro estimates of the productivity impact of health, which are trying to go around these problems.
- We will start by taking a step back and think about how to correctly estimate such effects.

The Rubin Causal Model

(Reference: Imbens and Woolridge, 2008).

- Consider a binary treatment W: 1 for treated, 0 for control, and an outcome Y (e.g. the treatment is : received an iron pill, the outcome could be: anemia, or earnings).
- Ex-ante, each individual i has two *potential outcomes*, $Y_i(1)$ if treated, $Y_i(0)$ if non-treated.

$$Y_i = Y_i(1)W_i + Y_i(0)(1 - W_i)$$

- The *treatment effect* for individual i is $Y_i(1) - Y_i(0)$.
- Ex-post, only one of the outcomes is realized: individual is treated or non-treated. Since no individual is observed both in the treated and non-treated state, we will not be able to estimate the treatment effect for each individual. All we can hope to estimate are some statistics concerning the treatment effect for a sample of individual.

Estimand

- We could be interested in the average treatment effect for the population: $E[Y_i(1) - Y_i(0)]$.
- we could want to know the average treatment effect for those who receive the treatment: $E[Y_i(1) - Y_i(0)|W_i = 1]$.
- Could be interested in the average treatment for those who have some characteristics (observed or unobserved): $E[Y_i(1) - Y_i(0)|X_i = x]$
- Or we may want to know other things about the treatment:
  - How the treatment is affecting the distribution in treatment and control groups (quantile treatment effects).
  - The quantile of treatment effects (this is not the same, and it is very hard to know!)

Estimating Average Treatment Effect

Suppose we have a population, with N_1 treated individual, and N_0 non treated individuals. Consider the difference between treated and control population:

$$E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 0]$$

$$= E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 1]$$
$$+E[Y_i(0)|W_i = 1] - E[Y_i(0)|W_i = 0]$$

$$= E[Y_i(1) - Y_i(0)|W_i = 1] + E[Y_i(0)|W_i = 1] - E[Y_i(0)|W_i = 0]$$

First term: ATT. Second term: difference in the underlying characteristics of the treated and non treated population (selection effect).

Selection mechanisms

Three cases:

- The probability of assignment does not depend on potential outcomes, and is a known function of covariates (*random assignment*). this case, $E[Y_i(0)|W_i = 1] = E[Y_i(0)|W_i = 0]$ and $E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 0]$ is an unbiased estimate of the effect of the treatment on the treated. Health example: Thomas et al (iron); Miguel and Kremer (worms)
- The probability of assignment does not depend on potential outcomes, but is an *unknown* function of covariates .

$$W_i \perp (Y_i(1), Y_i(0))|X_i$$

  (unconfoundness assumption, a.k.a. exogeneity, selection on observables). In this case, $E[Y_i(0)|W_i = 1, X = x] = E[Y_i(0)|W_i = 0, X = x]$, so the selection bias disappears if we appropriately control for x. *Matching, propensity score matching, regressions*, are various ways to deal with this.

Selection mechanisms (3)

- The probability of assignment depends on potential outcomes: there is a selection bias of unknown size. Program evaluation question is to find ways to deal with that. Leading strategies: *Difference-in-differences, Regression Discontinuity, Instrumental variables*.

Difference in Differences

Simplest setting:

- Individual i belong to one of groups $G = 1$, treated group, $G = 0$, non treated group.
- and is observed in one of two periods (or cohorts) $T = 1$ (post) and $T = 0$ (pre).
- Group $G = 1$ is treated when $T = 1$, not when $T = 0$.
- Identification Assumption: Potential outcome $Y_i(0)$ can be written:

$$Y_i(0) = \alpha + \beta T_i + \gamma G_i + \epsilon_i$$

  with $\epsilon_i \perp (T, G)$, i.e. ϵ_i is independent of the group indicator and its distribution does not change over time.
- Then: $Y_i(1) = Y_i(0) + \tau_{DID}$
- What is the key identification assumption?

Difference in difference estimator

$$\tau_{DID} = (E[Y_i|G=1,T=1) - E[Y_i|G=1,T=0])$$
$$-((E[Y_i|G=0,T=1) - E[Y_i|G=0,T=0]))$$

Sample equivalent:

- Replace expectation by population averages:

$$\tau_{DID} = (\overline{Y_{11}} - \overline{Y_{10}}) - (\overline{Y_{01}} - \overline{Y_{00}})$$

  where $\overline{Y_{gt}} = \frac{1}{N_{gt}} \sum_{G_i=g, T_i=t} Y_i$

- Or equivalently estimate OLS on

$$Y_i = \alpha_1 + \beta_1 T_i + \gamma_1 G_i + \tau_{DID}(T_i * G_i) + \epsilon_i$$

- Under the identification assumption, it is easy to show that τ_{DID} recovers the average treatment effect.

Example: Malaria Eradication in the Americas (Bleakley, 2007)

- Set-up:
  - Relatively swift ▸ Malaria Eradication Campaigns
  - Intensity of treatment depends on whether there was malaria before or not ▸ Figure
- Diff in Diff
  - Definition of treated and control cohorts in the US:
  - 1920 or later
  - Definition of treated and control regions
  - 1899 or earlier
  - Results: ▸ Regression
  - How would it look in a Diff and Diff table?
- Testing the identification assumption
  - Old versus very Old ▸ Results
  - Young versus very Young ▸ Results

Extension: Multiple groups, multiple Periods, or both

Let T denote the number of periods, and G the number of groups:

$$Y_i(0) = \alpha + \sum_{t=1}^{T} \beta_t 1[T_i = t] + \sum_{g=1}^{G} \gamma_g 1[G_i = g] + \epsilon_i$$

and $Y_i(1) = Y_i(0) + \tau_{DID}$
The model can be estimated with OLS regression:

$$Y_i = \alpha + \sum_{t=1}^{T} \beta_t 1[T_i = t] + \sum_{g=1}^{G} \gamma_g 1[G_i = g] + \tau_{DID} W_i + \epsilon_i$$

Where as before W_i is 1 for treated group for treated periods.

Extension: variable treatment intensity across periods

- Equivalent to have several treatments W_i^t, where W_i^t is equal to 1 for treated groups in year t

$$Y_i = \alpha + \sum_{t=1}^{T} \beta_t 1[T_i = t] + \sum_{g=1} \gamma 1[G_i = g] + \sum_{t=2}^{T} \tau_{tDID} W_i^t + \epsilon_i$$

(alternatively: compute a series of DID relative to one base period)

- Specification check: the treatment effect should follow the pattern of the extension of the program. It should be be 0 for all the periods before the treatment starts; it should equal for all periods where the treatment intensity was the same.
- In the malaria case, exposure depends on cohort of birth in a specific way: Exposition By Age
- We get this graph for the coefficients: encouraging?

Extension: Continuous treatment intensity across groups

- Suppose that the intensity of the treatment also depend on the group. We can think about this as if it were several treatments: $Y_i(w)$, for $w = 0, 1, 2, 3$.
- Alternatively, if we define $Wi = 1$ is the unit got any treatment, for some observable variable X, we may want to model $Y_i(Wi * X_i) = g(X_i) + Y_i(0)$
- For example, in Bleakley's case: X is the pre-campaign intensity level in the group, and he assumes linearity: $Y_i(g) = \tau_{CDID} M_g + Y(0)$
- With only two cohorts:

$$Y_i = \alpha + \beta T + \sum_{g=1}^{G} \gamma 1[G_i = g] + \tau_{CDID}(M_g * T_t) + \epsilon_i$$

▸ Bleakley, Table 2

Extension: Continuous treatment intensity across groups

- With more than 2 cohorts.
- Do the two cohorts approach cohort by Cohort, and graph the results: which pattern should it have? graph.
- He then tests whether the cohort effects have the right shape: Cohort pattern, US. Cohort pattern, Other countries
- Alternatively, we can follow the "multi-cohorts" approach:

$$Y_i = \alpha + \sum_{t=1}^{T} \beta_t^T 1[T_i = t] + \sum_{g=1}^{G} \gamma_g 1[G_i = g] + \sum_{t=2}^{T} \tau_{CDIDt}(M_g * T_t) + \epsilon_i$$

Table Appendix D.

Macro Implication

- Similar identification strategies are used by Cutler et al. (2008) and Lucas (2008) for looking at the impact of malaria on education. Results are quite comparable
- What are the macro-economic implications?
- Estimate that childhood infection decreases income by -0.5. Assuming no general equilibrium effect, this is also the GDP estimate
- Sachs's estimates translate into -2.16 (-1.3/0.6)
- This is much lower, but still significant (malaria would account for 10%-15% of the gap of Brazil and Mexico with the US)

Acemoglu and Johnson: A more macro approach

- A potential issue when going from micro-estimate to macro-estimate is the possibility of equilibrium effects: in the case of malaria, we compare cohorts. Maybe the younger cohorts are richer than the richer cohorts, but everybody is richer (or poorer) than they would be otherwise.
- The problem with macro setting is to find plausible source of variation
- Acemoglu and Johnson (2007) use the same identification strategy as Bleakley, but in a cross-country setting, for the disease against significant progress were made in the post-war period. (mainly turberculosis, pneumonia, malaria)
- Treatment intensity is a function of pre-campaign morbidity from those diseases (as in Bleakley).
- See graphs: Considerable gains in ▸ life expectancy and ▸ population.
- ... but ▸ no gains in GDP.
- so on balance a *loss* in GDP per capita.

Long Term Impact of Low Nutrition on Productivity

- "Barker" hypothesis (or fetal health). What matter in early childhood continue to matter later in life
- Evidence: Doblhammer–long term impact of month of birth, likely linked to nutrition available to mother.
- Almond, Qian: long term impact of famine in China (even on survivors, despite selection)
- Almond: people who were in gestation during 1918 influenza epidemics have lower life expectancy
- Banerjee, Duflo, Postel-Vinay and Watts: impact of shock at birth on height at 20.
- Field: Iodine supplementation

Iodine supplementation in Tanzania

- 1 billion people at risk of iodine shortage (old soil and no seafood)
- Iodine deficiency in first trimester of pregnancy thought to lead to permanent irreversible brain damage, apparently especially for girls.
- Tanzania had an intensive campaign to distribute iodine capsules, ultimately reaching 25% of the population, starting in 1986, and targeted to the 25 districts that had the largest goiter rates.
- In principle, women must receive a capsule every 2 years (duration of the dose). In practice some districts started later, and the distribution was not every 2 years: table 1

Strategy and results

- Authors calculate the probability that a child born in a given month was covered when in Utero (as a function of when the pills were distributed in the district) and introduce district fixed effect and month fixed effect (compare children born at the wrong time in treated district).
- They also carry out the analysis within households (siblings born a little bit too late or too early).
- Results in table 3 table 4 : Large effects, especially for girls.
- Robustness: Given the dose, effect should be highest when IDD is not too high and not too low: compare results across regions which produce more or less cassava: table 6
- Other than the effect through cognition, what could be the channel through which this intervention affect education? What regression can they run to rule them out?

Conclusion: Back to Das Gupta and Ray

- There is a strong relationship between health and productivity at the micro level (and also between education and productivity)
- Role of Micro-nutrients seems particularly important (iodine, iron: Thomas).
- No very solid estimate of the impact of nutrition on productivity (nobody does that!) but earlier estimates suggest an elasticy of about 0.4 (Strauss).
- Impact of nutrition in-utero and in childhood may be much larger than later in life, since it may cause permanent damage on health (so impact would be multiplied by years of life), and also through amplification impacts through education.
- Need to go back to thinking in more detail about what is happening within the household: if nutrients are indeed shared more unequally in the household when there is a shock (as the Das Gupta Ray model would suggest), this may create a space for a inter-generational poverty trap to emerge.

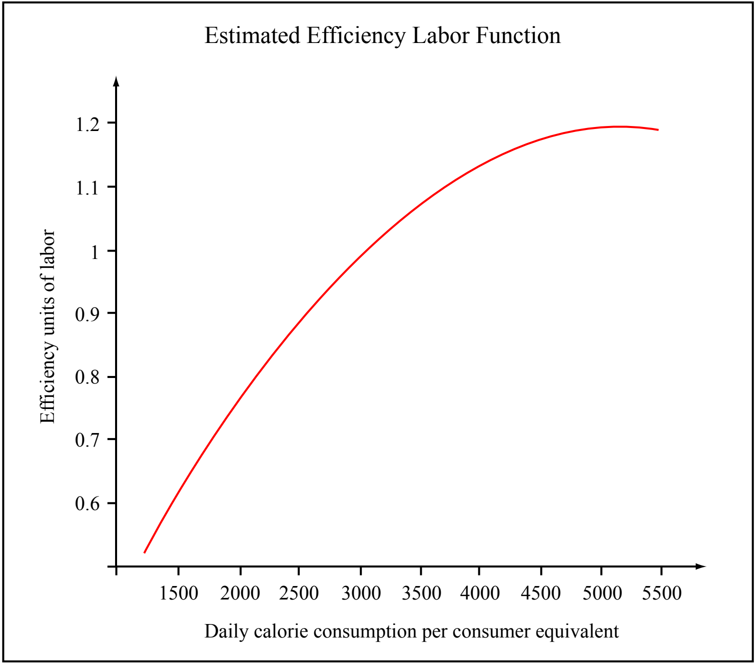
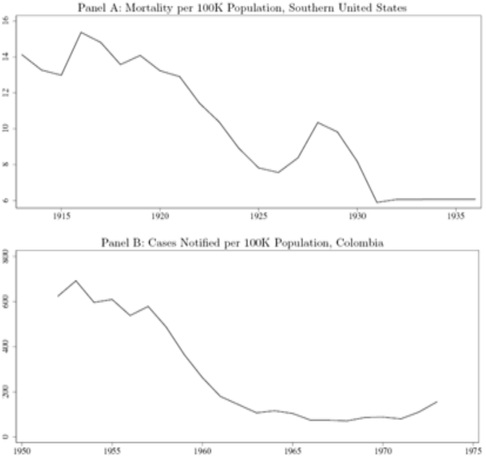


Figure by MIT OpenCourseWare.

Bleakley (2006)

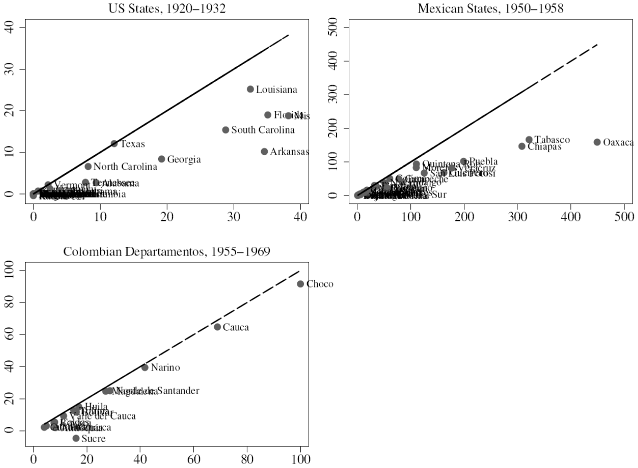
Figure 1: Malaria Incidence Before and After the Eradication Campaigns



Notes: Panel A plots the estimated malaria mortality per capita for the Southern region and bordering states. Because the death registration system was being phased in over the period, a regression model with state fixed effects is used to control for sample changes, and the time series is constructed from the year dummies in the regression, normalized to match the end-of-period data when all states were represented. (Census Bureau *Vital Statistics*, various years, and author's calculations.) Panel B reports data on notified cases of malaria for Colombia (SEM, 1979).

Bleakley (2006)

Figure 2: Highly Infected Areas Saw Greater Declines in Malaria



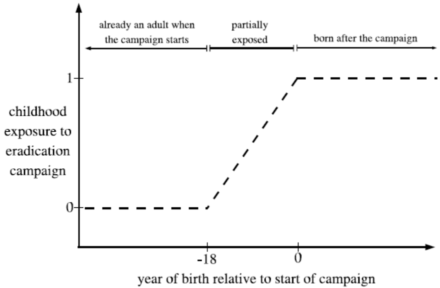
Notes: The y axis displays the estimated decrease in malaria mortality post-intervention. The x axis is the pre-campaign malaria mortality rate. The 45-degree line represents complete eradication. Both variables are expressed per 100,000 population. United States data are reported in Maxcy (1923) and *Vital Statistics* (Census, 1933). Mexican data are drawn from Pesqueira (1957) and from the Mexican *Anuario Estadístico* (Dirección General de Estadística, 1960). SEM (1957) and the Colombian *Anuario de Salubridad* (DANE, 1968-70) are the sources for the Colombian data.

```
.
. **************** PART 2 *****************
. g byte mhigh_young= mal_high*young

. reg sei mhigh_young mal_high young if (young==1 | old==1) & mal_highlow, r

Linear regression                                      Number of obs =    2044
                                                       F(  3,  2040) =  501.00
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.3374
                                                       Root MSE      =  .30635

------------------------------------------------------------------------------
             |               Robust
         sei |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 mhigh_young |    .159573   .0231584     6.89   0.000     .1141563    .2049897
    mal_high |  -.2190388   .0200889   -10.90   0.000    -.2584357    -.179642
       young |   .3299299    .014544    22.68   0.000     .3014073    .3584525
       _cons |   3.375146   .0133338   253.13   0.000     3.348997    3.401295
------------------------------------------------------------------------------
```

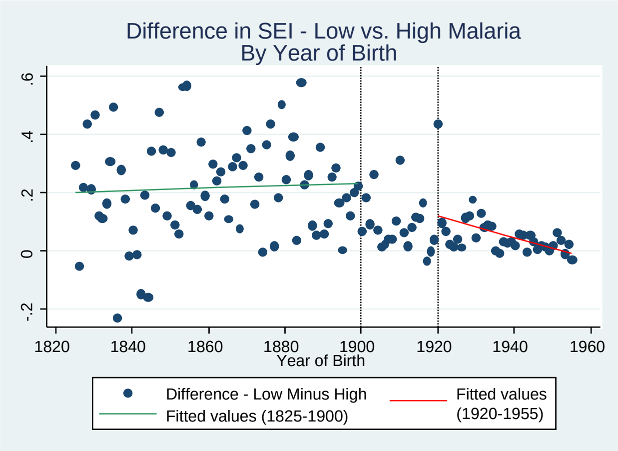
```
.
. reg sei mal_high yold mhigh_yold if old & mal_highlow, r

Linear regression                                      Number of obs =    1279
                                                       F(  3,  1275) =   58.19
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.1257
                                                       Root MSE      =  .35491

------------------------------------------------------------------------------
             |               Robust
         sei |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    mal_high |  -.2069026   .0321059    -6.44   0.000    -.2698887   -.1439165
        yold |   .1692435   .0259945     6.51   0.000     .1182468    .2202402
  mhigh_yold |  -.0152891   .0405002    -0.38   0.706    -.0947435    .0641654
       _cons |   3.268976   .0197418   165.59   0.000     3.230246    3.307706
------------------------------------------------------------------------------

. reg sei mal_high vyoung mhigh_vyoung if young & mal_highlow, r

Linear regression                                      Number of obs =     765
                                                       F(  3,   761) =   19.75
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0729
                                                       Root MSE      =  .16983

------------------------------------------------------------------------------
             |               Robust
         sei |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    mal_high |  -.0775506   .0165237    -4.69   0.000    -.1099882   -.0451131
      vyoung |   .0440054   .0107842     4.08   0.000     .0228351    .0651756
mhigh_vyoung |   .0519633    .018939     2.74   0.006     .0147845    .0891421
       _cons |   3.689949   .0077427   476.57   0.000      3.67475    3.705149
------------------------------------------------------------------------------
```

Bleakley (2006)

Figure 3: Childhood Exposure to Eradication Campaign



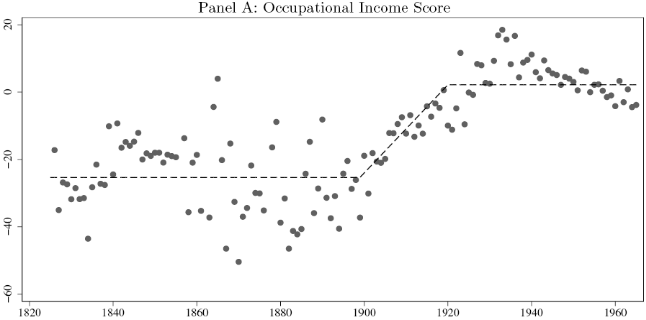
Notes: This graph displays on the fraction of childhood that is exposed to a hypothetical (and instantaneous) campaign as a function of year of birth minus the start year of the campaign.



Difference in SEI - Low vs. High Malaria
By Year of Birth
-.2
0
.2
.4
.6
1820
1840
1860
1880
1900
1920
1940
1960
Year of Birth
Difference - Low Minus High
Fitted values (1825-1900)
Fitted values (1920-1955)

Bleakley (2006)

Figure 4: Cohort-Specific Relationship: States in the U.S.



Notes: These graphics summarize regressions of income proxies on pre-eradication malaria-mortality rates (measured by the Census in 1890). The y axis for each graphic plots the estimated cohort-specific coefficients on the state-level malaria measure. The x axis is the cohort's year of birth. Each cohort's point estimate is marked with a dot. The dashed lines measure the approximate number of years of potential childhood exposure to the malaria-eradication activities in the South. For each year-of-birth cohort, OLS regressions coefficients are estimated on the cross section of states of birth. The state-of-birth average outcome is regressed onto malaria, Lebergott's (1964) measure of 1899 wage levels, a dummy for the Southern region, and the various control variables described in Appendix C. Appendices A and B describe, respectively, the outcome variables and the malaria measure.

Bleakley (2006)

Table 1: Exposure to Malaria Eradication versus Trends

Degree of Polynomial-Trend Control:	0	1	2	3
Outcome Variables:				
	Panel A: United States			
Occupational Income Score	28.684 ***	33.802 ***	34.611 ***	34.235 ***
	(1.509)	(3.664)	(4.105)	(5.412)
	{0.109}	{0.129}	{0.132}	{0.130}
Duncan's Socioeconomic Index	52.549 ***	48.862 ***	57.078 ***	55.248 ***
	(2.956)	(6.654)	(7.485)	(9.782)
	{0.158}	{0.147}	{0.172}	{0.166}
	Panel B: Brazil			
Literacy	0.029 ***	0.018 ***	0.017 ***	0.002
	(0.002)	(0.004)	(0.004)	(0.006)
	{0.152}	{0.094}	{0.089}	{0.010}
Years of Schooling	0.214 ***	0.116 *	0.349 ***	0.179 **
	(0.025)	(0.070)	(0.057)	(0.090)
	{1.120}	{0.607}	{1.827}	{0.937}
Log Total Income	0.073 ***	0.094 ***	0.104 ***	0.084 ***
	(0.005)	(0.011)	(0.011)	(0.019)
	{0.382}	{0.492}	{0.544}	{0.440}
Log Earned Income	0.056 ***	0.080 ***	0.082 ***	0.048
	(0.008)	(0.022)	(0.025)	(0.054)
	{0.293}	{0.419}	{0.429}	{0.251}

Bleakley (2006)

Table 1: Exposure to Malaria Eradication versus Trends

Degree of Polynomial-Trend Control:	0	1	2	3
Outcome Variables:				
	Panel C: Colombia			
	0.023 **	0.047 *	0.058 **	-0.019
Literacy	(0.011)	(0.026)	(0.028)	(0.052)
	{0.009}	{0.019}	{0.023}	{-0.008}
	0.800 ***	0.854 **	0.683 **	0.673
Years of Schooling	(0.131)	(0.358)	(0.340)	(0.601)
	{0.317}	{0.338}	{0.270}	{0.267}
	0.170 ***	0.104 **	0.121 ***	-0.146
Industrial Income Score	(0.016)	(0.047)	(0.044)	(0.090)
	{0.067}	{0.041}	{0.048}	-{0.058}
	Panel D: Mexico			
	0.008 ***	-0.006	-0.009 *	0.008
Literacy	(0.003)	(0.004)	(0.005)	(0.007)
	{0.026}	{-0.019}	{-0.029}	{0.026}
	-0.087 ***	-0.194 ***	-0.178 ***	-0.021
Years of Schooling	(0.020)	(0.051)	(0.046)	(0.077)
	{-0.279}	{-0.623}	{-0.571}	{-0.067}
	0.067 ***	0.021	0.063 **	-0.050
Log Earned Income	(0.016)	(0.035)	(0.026)	(0.070)
	{0.215}	{0.067}	{0.202}	{-0.160}

Notes: This table reports estimates of the childhood-exposure variable in equation 2 using OLS. The outcome variables used to construct the time series of $\hat{\beta}_k$ are as indicated in each row. Robust (Huber-White) standard errors in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Observations are weighted by the inverse of the coefficient's standard error. Reporting of additional terms suppressed. The terms in curly brackets report the point estimate multiplied by the difference between 95th and 5th percentile malaria intensity. For the United States, this number is also normalized by the average value of the relevant income proxy for white males born in the South between 1875

Bleakley (2006)

Appendix D: Panel Estimates of Childhood Exposure

Dependent Variables: Income Measures by Cohort

Panel A: United States

Occupational Income Score	Duncan Index
27.936 ***	52.040 ***
(5.528)	(13.176)
{0.106}	{0.157}

Panel B: Brazil

Literacy	Education	Log Total Income	Log Earned Income
0.030 ***	0.171 ***	0.070 ***	0.046 ***
(0.005)	(0.051)	(0.021)	(0.017)
{0.156}	{0.894}	{0.366}	{0.242}

Panel C: Colombia

Literacy	Education	Income Index
0.209 ***	1.751 ***	0.149 ***
(0.032)	(0.425)	(0.051)
{0.083}	{0.694}	{0.059}

Panel D: Mexico

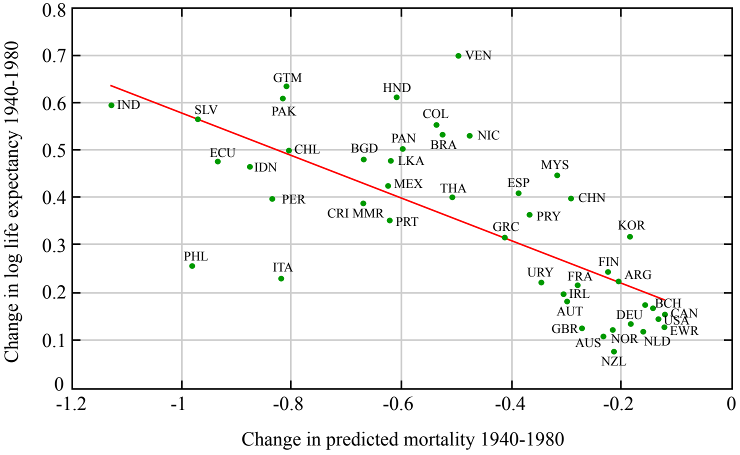
Literacy	Education	Log Earned Income
0.047 ***	0.229 *	0.151 ***
(0.011)	(0.140)	(0.036)
{0.150}	{0.736}	{0.486}

Bleakley (2006)

Table 2: Cross-Cohort Differences and Malaria: United States

	Malaria Mortality (Fraction of Total), 1890		Malaria Ecology (Mellinger)		Malaria Ecology (Hong)		Malaria Mortality (per 100K Population), 1920	
Dependent Variable:								
Occupational Income Score	X		X		X		X	
Duncan's Socioeconomic Index		X		X		X		X
Specification:	*Panel A: Basic Results*							
OLS, Basic Specification	37.927 ***	60.316 ***	0.570 **	1.191 **	16.278 ***	19.608 ***	0.030	0.057
	(11.101)	(21.311)	(0.267)	(0.535)	(2.040)	(4.737)	(0.021)	(0.036)
	{0.144}	{0.182}	{0.032}	{0.052}	{0.265}	{0.253}	{0.050}	{0.074}
2SLS, Using the Other Three Proxies as Instruments	44.367 ***	71.573 ***	1.312 *	2.064 *	15.133 ***	23.345 ***	0.074 **	0.110 **
	(14.238)	(24.199)	(0.748)	(1.075)	(3.813)	(8.205)	(0.031)	(0.053)
	{0.169}	{0.216}	{0.073}	{0.091}	{0.247}	{0.301}	{0.122}	{0.144}
Additional Controls:	*Panel B: Alternative Control Sets*							
Health	33.897 ***	63.480 ***	0.483 ***	1.078 ***	15.171 ***	24.580 ***	0.038	0.066
	(9.733)	(20.610)	(0.183)	(0.346)	(2.506)	(5.092)	(0.025)	(0.044)
Education	44.825 ***	59.306 **	0.552 **	1.080 ***	14.119 ***	16.543 ***	0.062 **	0.063
	(12.240)	(23.279)	(0.268)	(0.412)	(2.093)	(5.067)	(0.024)	(0.046)
Other	30.118 ***	45.827 **	0.388 **	1.050 ***	12.423 ***	13.082 ***	0.029	0.006
	(11.400)	(18.134)	(0.162)	(0.367)	(2.083)	(4.751)	(0.038)	(0.045)
Full Controls	33.392 **	59.257 **	0.385	0.985 **	15.564 ***	24.357 ***	0.048	0.060
	(13.844)	(29.103)	(0.236)	(0.473)	(3.280)	(7.088)	(0.030)	(0.056)

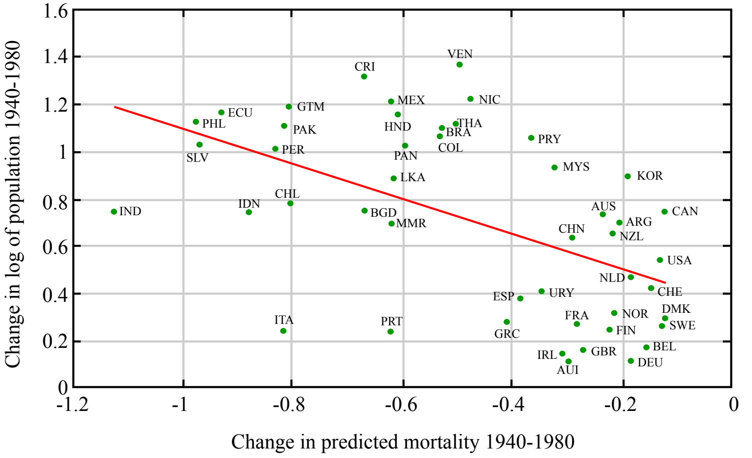
Notes: This table reports estimates of equation 3 using OLS and 2SLS. The units of observation are U.S. states. The dependent variables are as indicated in the column headings. Robust (Huber-White) standard errors in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Reporting of constant term suppressed. Unexposed cohorts are those born before 1890 and fully exposed cohorts are those born after 1920. Cohorts are determined based on state of birth. The universe for the base sample consists of the native-born white population between the ages of 25 and 55 (15–55 for literacy) in the 1880–2000 census microdata from the IPUMS and NAPP databases. The terms in curly brackets report the point estimate multiplied by the difference between 95th and 5th percentile malaria intensity and normalized by the average value of the relevant income proxy for white males born in the South between 1875 and 1895. The specification for the basic results includes the malaria variable, a dummy for Southern birthplace, and the Lebergott (1964) measure of average unskilled wage in the state of birth. Appendices A and B describe, respectively, the outcome variables and malaria measures. The additional controls are described in the text and Appendix C.



Change in log life expectancy and change in predicted mortality, 1940-80, base sample.

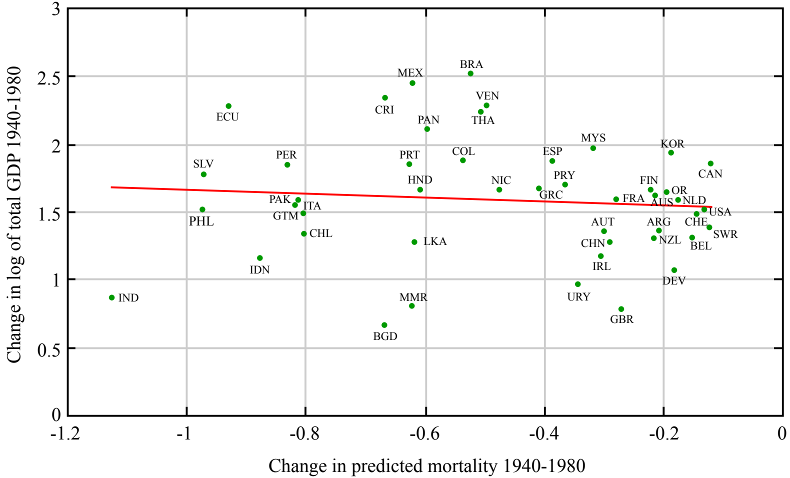
Figure by MIT OpenCourseWare.

Acemoglu and Robinson (2007)



Change in log of population and change in predicted mortality, 1940-80, base sample.

Figure by MIT OpenCourseWare.



Change in log of total GDP and change in predicted mortality, 1940-80, base sample.

Figure by MIT OpenCourseWare.

Field, Robles, and Torero (2007)

Table 1: Summary of Timing and Coverage of Intervention Across Districts

	Region	District	Year of Intervention (Coverage - %)* 1	2	3	4	5	Average Frequency (yr)
1	Dodoma	Mpwapwa	1990 (65)	1992 (58)				2.00
2	Arusha	Monduli	1992 (71)					n/a
3	Arusha	Arumeru	1991 (89)					n/a
4	Kilimanjaro	Rombo	1990 (68)					n/a
5	Morogoro	Ulanga	1988 (73)	1991 (61)	1992 (34)			1.33
6	Ruvuma	Songea Rural	1987 (91)	1991 (74)	1995 (85)			2.67
7*	Ruvuma	Mbinga	1995 (92)					n/a
8	Iringa	Mufindi	1986 (41)	1991 (63)	1995 (54)			3.00
9	Iringa	Makete	1986 (20)	1991 (62)	1993 (62)	1996 (49)		2.50
10	Iringa	Njombe	1989 (76)	1992 (68)	1995 (64)			2.00
11	Iringa	Ludewa	1989 (59)	1992 (62)	1995 (47)			2.00
12	Mbeya	Chunya	1990 (49)					n/a
13	Mbeya	Mbeya Rural	1986 (44)	1989 (84)	1990 (90)	1993 (53)	1997 (53)	1.75
14	Mbeya	Kyela	1989 (91)	1993 (57)				4.00
15	Mbeya	Rungwe	1986 (35)	1990 (73)	1993 (49)			2.33
16	Mbeya	Ileje	1989 (94)	1992 (71)				3.00
17	Mbeya	Mbozi	1989 (67)	1991 (63)				2.00
18	Rukwa	Mpanda	1987 (79)	1991 (60)	1993 (72)			2.00

Field, Robles, and Torero (2007)

Table 3. Grade Attainment and IOC Supplementation in Utero

	Boys and girls	Boys	Girls	Boys and girls	Boys	Girls	Boys	Girls
IOC in utero	0.357	0.315	0.77	0.157	-0.027	0.383		
(IOC in utero =born 1-3 years after program)	[0.142]*	[0.267]	[0.298]*	[0.112]	[0.154]	[0.166]*		
IOC in utero							-0.023	0.408
(IOC in utero =born 1-2 years after program)							[0.177]	[0.187]*
Age 11	0.558	0.571	0.743	0.401	0.45	0.296	0.451	0.251
	[0.140]**	[0.272]*	[0.256]**	[0.091]**	[0.125]**	[0.135]*	[0.126]**	[0.135]
Age 12	1.293	1.237	1.531	1.18	1.206	1.148	1.209	1.08
	[0.118]**	[0.216]**	[0.234]**	[0.086]**	[0.120]**	[0.124]**	[0.121]**	[0.125]**
Age 13	2.049	1.952	2.657	1.866	1.714	2.015	1.719	1.941
	[0.148]**	[0.278]**	[0.293]**	[0.096]**	[0.132]**	[0.141]**	[0.127]**	[0.134]**
Female	0.247			0.213				
	[0.090]**			[0.063]**				
Mother < age 23 at birth	-0.195	0.128	-0.31	0.071	0.022	0.11	0.024	0.093
	[0.202]	[0.356]	[0.501]	[0.070]	[0.096]	[0.103]	[0.095]	[0.101]
Number same sex siblings				0.208	0.187	0.25	0.187	0.258
				[0.076]**	[0.104]	[0.113]*	[0.104]	[0.113]*
Fixed effects	*House-hold*	*House-hold*	*House-hold*	*District*	*District*	*District*	*District*	*District*
Observations	*2251*	*1154*	*1097*	*2251*	*1154*	*1097*	*1154*	*1097*

Notes: Data from the 2000 Tanzanian Household Budget Survey, sample restricted to children ages 10-13 in 27 districts targetted for iodized oil capsule (IOC) distribution between 1986 and 1995. In columns 1-6, IOC in utero is equal to 1 or 0.5 (depending on mother's age at birth) if a child was born 1-3 years after IOC was distributed in the district; in columns 7-8, IOC in utero is equal to 1 or 0.5 (depending on mother's age at birth) if a child was born 1-2 years after IOC was distributed in the district. All regressions control for birth order and sex-specific birth order. * significant at 5%; ** significant at 1%

Field, Robles, and Torero (2007)

Table 4: Grade Attainment and IOC Supplementation in Utero, 2004

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Grade attainment, ages 10-14			*Grade attainment, ages 10-12*			*Enter secondary school, ages 10-14*		
	All	Girls	Boys	All	Girls	Boys	All	Girls	Boys
IOC in utero	0.487	0.870	0.403	0.189	1.604	0.276	0.081	0.139	0.096
	[0.184]**	[0.424]*	[0.331]	[0.150]	[0.606]*	[0.662]	[0.031]**	[0.081]+	[0.058]+
Age 11	0.317	0.302	0.405	-0.191	-0.593	0.585	-0.016	-0.045	0.02
	[0.108]**	[0.247]	[0.195]*	[0.098]+	[0.454]	[0.420]	[0.018]	[0.047]	[0.034]
Age 12	0.94	0.884	1.135	-0.124	0.644	1.203	-0.001	-0.098	0.065
	[0.124]**	[0.300]**	[0.226]**	[0.086]	[0.547]	[0.441]**	[0.021]	[0.058]+	[0.040]
Age 13	1.349	1.682	1.267				0.035	-0.03	0.102
	[0.166]**	[0.428]**	[0.307]**				[0.028]	[0.082]	[0.054]+
Age 14	2.036	2.185	2.152				0.117	0.005	0.234
	[0.202]**	[0.507]**	[0.375]**				[0.034]**	[0.097]	[0.066]**
Month of birth	-0.027	-0.051	-0.024	0.003	-0.030	0.018	-0.002	-0.005	-0.004
	[0.010]**	[0.023]*	[0.017]	[0.009]	[0.031]	[0.029]	[0.002]	[0.004]	[0.003]
Female	0.352			0.309			0.015		
	[0.062]**			[0.060]**			[0.010]		
Household fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	*3672*	*1797*	*1875*	*4984*	*1147*	*1178*	*3672*	*1797*	*1875*

Notes: Data from the 2004 Tanzanian Demographic and Health Survey, sample restricted to children ages 10-14. IOC in utero is equal to (birth month/12) if a child was born 1-3 years after IOC was distributed in the district. All regressions control for dummy indicators of birth order and sex-specific birth order. + significant at 10%; * significant at 5%; ** significant at 1%.

Field, Robles, and Torero (2007)

Table 6: Variation in Effect on Schooling of IOC Supplementation in Utero

	Rate of Cassava Consumption in District			*Amount of IOC*	
	High (0.41-0.62)	Medium (0.10-0.40)	Low (< 0.10)	Mother>22 at birth (380 mg)	Mother<23 at birth (200 mg)
IOC in utero	0.046	0.508	-0.02	0.431	0.066
(IOC in utero =born 1-3 years after program)	(0.391)	(0.165)**	(0.252)	(0.198)*	(0.199)
Female	0.417	0.172	0.154	0.252	0.304
	(0.188)*	(0.146)	(0.148)	(0.146)	(0.156)
Age 11	0.68	0.46	0.451	0.783	0.401
	(0.281)*	(0.222)*	(0.245)	(0.225)**	(0.242)
Age 12	1.64	1.224	0.967	1.524	1.232
	(0.241)**	(0.179)**	(0.205)**	(0.195)**	(0.210)**
Age 13	2.086	2.051	1.724	2.27	1.684
	(0.298)**	(0.232)**	(0.295)**	(0.253)**	(0.263)**
Household fixed effects	yes	yes	yes	yes	yes
Observations	*669*	*804*	*778*	*983*	*799*

Notes: Data from the 2000 Tanzanian Household Budget Survey, sample restricted to children ages 10-13 in 27 districts targetted for iodized oil capsule (IOC) distribution between 1986 and 1995. Children and women below age 23 were given IOC containing 200mg of iodine and women over 22 were given IOC containing 380 mg of iodine. In all regressions, IOC in utero is equal to one if a child was born 1-3 years after IOC was distributed in the district. Regressions also control for birth order and sex-specific birth order. Rate of cassava consumption defined as fraction of THBS households in district that report growing cassava.

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

Private and Social Returns to Education

Esther Duflo

14.771

Education and Development

- Tremendous correlation between education and level of income across countries.
- The R2 of the regression in figure 1 is 0.65. Human capital is given a weight of two thirds in Cobb Douglas models.
- Cross Countries studies often regress GDP *growth* on level of education, and also find large coefficient (one extra year of average schooling is associated with 0.3 percent extra growth every year in GDP, between 1960 and 1990).
- This raises a number of questions:
    - Sources of this strong correlation (in level and in the growth regressions)
    - If education is so important, need to understand the determinants of its provision, who should pay for it, the optimal way to pay for it, etc.

Mincerian Returns to Education

- Mincer hypothesizes that each extra year of education raise income by b%.

$$y_i = a + bS_i + cE_i + \epsilon_i$$

- Where S is schooling and E is experience.
- Why call this returns to education?
- Social returns may differ from private returns:
  - Costs
  - Externalities

Estimating Returns to Education

- This question can be estimated from micro data.
- Concerns:
  - Functional form (why log linear? Convex? Concave?)
  - Omitted variables
- Randomly identifying education is not easy, and convincing control strategies are difficult to come by.
- Therefore a large literature in labor searches for *instruments*: something that affects educational achievement but does not affect income directly.

Instrumental Variables

- Let Z_i be an *instrument*, which affects the probability that an individual is treated
- Let $W_i(1)$ be the treatment status for individual i if $Z = 1$, and $W_i(0)$ the treatment status of the same individual if $Z_i = 0$.
- The observed treatment is : $W_i = Z_i W_i(1) + (1 - Z_i) W_i(0)$
- As before, $Y_i(1)$ is potential outcome of treated (if $W_i = 1$) and $Y_i(0)$ is potential outcome if non-treated.
- Identification assumptions (Imbens and Angrist):
  1. *All* Potential outcomes are independent of the Instrument

     $$(Y_i(1), Y_i(0), W_i(1), W_i(0)) \perp Z_i$$

  2. What does this imply?
     - Treatment assignment is randomly assigned (or can be treated as such)
     - Treatment has no direct impact on the outcome (that is not implied by randomization of the instrument and has to be argued on a case by case basis!)
  3. Monotonicity: $W_i(1) \geq W_i(0)$ for everyone

More on Monotonicity

Three groups of people :

1. The Compliers: $Y_i(1) = 1$ and $Y_i(0) = 0$.
2. The Never-Takers: $Y_i(1) = 0$ and $Y_i(0) = 0$
3. The Always-Takers: $Y_i(1) = 1$ and $Y_i(0) = 1$
4. The Defiers: $Y_i(1) = 0$ and $Y_i(0) = 1$

The monotonicity assumption means that there are no defiers.
This is not a testable assumption, and needs to be assessed on a case by case basis.

Wald estimate and its interpretation

Wald estimate: Ratio of *Reduced form* and *First stage*.

$$\hat{\beta}_{IV} = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[W_i|Z_i = 1] - E[W_i|Z_i = 0]}$$

Case of constant treatment effect:

$$Y_i = a + bW_i + \epsilon_i$$

$$W_i = \alpha + \gamma Z_i + \upsilon_i$$

Substituting:

$$Y_i = a + b(\alpha + \gamma Z_i + \upsilon_i) + \epsilon_i$$

$$Y_i = a + \pi Z_i + \omega_i$$

Independence assumption insures that $\omega_i = \epsilon_i + b\upsilon_i$

$$b = \frac{\pi}{\gamma}$$

Two stage least squares

- Regress W on Z
- Regress Y on predicted W
- (in practice this is done in one step by the "two stage least square" procedure)
- Can be generalize to multiple instruments (and multiple treatment):
  1. Project (regress) X onto the vector of instruments Z
  2. Regress Y on the predicted value of X

$$\beta_{2SLS} = (W'Z(Z'Z)^{-1}Z'W)W'Z(Z'Z)^{-1}Z'Y$$

- Intuition: we are only using the part of the variance in the X for which we believe the identification assumptions.

Heterogenous treatment Effect

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$$

$$= E[W_i(1)Y_i(1) + (1 - W_i(1))Y_i(0)|Z_i = 1]$$
$$-E[W_i(0)Y_i(1) + (1 - W_i(0))Y_i(0)|Z_i = 0]$$

$$= E[(W_i(1) - W_i(0))(Y_i(1) - Y_i(0))] + E[Y_i(0)|Z_i = 1] - E[Y_i(0)|Z_i = 0]$$

$= E[(W_i(1) - W_i(0))(Y_i(1) - Y_i(0))]$ (by independence)
$= E[-(Y_i(1) - Y_i(0))|W_i(1) - W_i(0) = -1]P(W_i(1) - W_i(0) = -1)$
$+E[0 * (Y_i(1) - Y_i(0))|W_i(1) - W_i(0) = 0]P(W_i(1) - W_i(0) = 0)$
$+E[(Y_i(1) - Y_i(0))|W_i(1) - W_i(0) = 1]P(W_i(1) - W_i(0) = 1)$
$= E[Y_i(1) - Y_i(0)|W_i(1) - W_i(0) = 1] * P(W_i(1) - W_i(0) = 1)$
(by monotonicity)
$= E[Y_i(1) - Y_i(0)|W_i(1) - W_i(0) = 1] * (E[W_i(1)] - E[W_i(0)]$

Wald Estimate is treatment effect on the compliers

$$\hat{\beta}_{IV} = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[W_i|Z_i = 1] - E[W_i|Z_i = 0]}$$

$$= E[Y_i(1) - Y_i(0)|W_i(1) - W_i(0) = 1]$$

Who are the compliers?

- Special case: Treatment on the Treated:
  - When $W_i(0) = 0$ (e.g. randomized evaluation: all the control stays control)
- General case: Those are compelled by the instrument to get the treatement: external validity?
- While we cannot know who the compliers are, we can describe their characteristics

The INPRES Experiment: First Stage and Reduced form

- The Set up is a DID set-up similar to Bleakley's: Cohorts and Region
  - School construction campaign started in 1973: affect cohort age 12 or younger in 1973
  - More schools were built in regions that were initially lagging behind in term of education
- Results: Impacts of the program on Education and on log(wages)
  - ▸ Basic DID
  - ▸ Placebo experiment
  - Use all the regional variation (keep 2 cohorts) ▸ Table
  - Use all the regional variation, and all cohorts
- Check identification assumptions by estimating effects for all the cohorts ▸ Graph
- Force the earlier cohort to have an 0 effect: more precision ▸ Table

Instrumental variable

- What can we use as instruments?
  - If we wanted to use just one instrument
  - If we wanted to use many instruments?
- What are the identification assumptions? Do we believe in them?
- ▸ Results Did the IV make a big difference?
- What is the interpretation of the estimate? What are the years of education we are estimating the returns for?
- Interpretation of IV when the treatment takes more than one value: weighted average of marginal effects (going from 0 to 1, 1 to 2, etc..), where the weights are the fraction of people who are moved from one value of the instrument to another.
- See ▸ impact of programs by year of education

Reconciling Macro and Micro pictures

- Returns to education estimated from Mincerian specification ranges from 2.7% to 15.4%. Mean of 9%, stdv 2.2%. Generally at individual level IV is roughly equal to OLS.
- Puzzle 1: levels
- Countries in top decile of education distribution have about 8 more years of education than those in the bottom. They should have GDP no more than twice the size if private returns were the only part of the story. In fact they are about 15 times richer.
- Puzzle 2: Does the effect of *level* of education on *Growth* of GDP follows from the Mincer Framework? What can explain an effect of *level* of education on *growth* of income.
- Potential solution to both puzzle: Externalities.

Estimating Externalities

- The same experiment can be used to estimate the "social returns" to education
- Do we expect externalities to be positive or negative? (why?)
- We are looking to estimate:

$$y_i = \alpha + \beta S_i + \beta \overline{S_i} + \epsilon_i$$

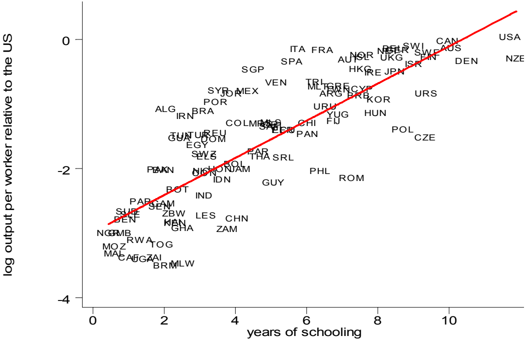
- Two estimation problem: we need an instrument for S_i and an instrument for $\overline{S_i}$ (Acemoglu and Angrist).
- Consider a cohort who was 12 or older in 1973, and is thus not exposed by the program
- Until 1979, no-one in the labor market is educated in the new schools.
- Starting in 1979, slow influx of the graduate of the new schools ▸ Graph

Empirical Strategy

- Fix the cohort, let the years vary.
- Survey Year*Region are instrument for $\overline{S_i}$. Are they correlated with S_i?
- Results (▸ Graph, ▸ Table): Mushy, but if anything, equilibrium effects are negative.

Reconciling Macro and Micro picture (2)

- Externalities are not doing the trick...
- Other potential explanations:
  - Omitted variable
  - Endogeneity: Future growth in income motivates people to invest in education (Bills and Klenow)
- Micro-evidence of this channel
  - Foster+Rosenzweig HYV revolution in India (AER, 1995)
  - Jensen, Nguyen: Young people sensitive to *perceived* returns to education.


log output per worker relative to the US
years of schooling
0
-2
-4
0
2
4
6
8
10
USA
CAN
AUS
SWI
DEN
NZE
ITA
FRA
SPA
SGP
UKG
ISR
JPN
IRE
VEN
URS
KOR
HUN
POL
CZE
MEX
POR
ALG
IRN
BRA
YUG
FIJ
URU
CHI
PAN
EGY
DOM
THA
SRL
PHL
ROM
GUY
IDN
IND
LES
CHN
ZAM
TOG
RWA
MOZ
MLW
ZBW
GHA

Duflo (2001)

Means of Education and Log(Wage) by Cohort and Level of Program Cells

	Years of education			*Log(wages)*		
	Level of program in region of birth			*Level of program in region of birth*		
	High (1)	*Low (2)*	*Difference (3)*	*High (4)*	*Low (5)*	*Difference (6)*
Panel A: Experiment of interest						
Aged 2 to 6 in 1974	8.49 (0.043)	9.76 (0.037)	-1.27 (0.057)	6.61 (0.0078)	6.73 (0.0064)	-0.12 (0.010)
Aged 12 to 17 in 1974	8.02 (0.053)	9.40 (0.042)	-1.39 (0.067)	6.87 (0.0085)	7.02 (0.0069)	-0.15 (0.011)
Difference	0.47 (0.070)	0.36 (0.038)	0.12 (0.089)	-0.26 (0.011)	-0.29 (0.0096)	0.026 (0.015)
Panel B: Control experiment						
Aged 12 to 17 in 1974	8.02 (0.053)	9.40 (0.042)	-1.39 (0.067)	6.87 (0.0085)	7.02 (0.0069)	-0.15 (0.011)
Aged 18 to 24 in 1974	7.70 (0.059)	9.12 (0.044)	-1.42 (0.072)	6.92 (0.0097)	7.08 (0.0076)	-0.16 (0.012)
Difference	0.32 (0.080)	0.28 (0.061)	0.034 (0.098)	0.056 (0.013)	0.063 (0.010)	0.0070 (0.016)

Notes: The sample is made of the individuals who earn a wage. Standard errors are in parentheses.

Figure by MIT OpenCourseWare.

Effect of the Program on Education and Wages: Coefficients of the Interactions Between Cohort Dummies and the Number of Schools Constructed per 1,000 Children in the Region of Birth

		Dependent variable					
		Years of education			*Log(hourly wage)*		
	Observations	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Panel A: Experiment of interest: Individuals aged 2 to 6 or 12 to 17 in 1974 (Youngest cohort: Individuals ages 2 to 6 in 1974)							
Whole sample	78,470	0.124 (0.0250)	0.15 (0.0260)	0.188 (0.0289)			
Sample of wage earners	31,061	0.196 (0.0424)	0.199 (0.0429)	0.259 (0.0499)	0.0147 (0.00729)	0.0172 (0.00737)	0.0270 (0.00850)
Panel B: Control Experiment: Individuals aged 12 to 24 in 1974 (Youngest cohort: Individuals ages 12 to 17 in 1974)							
Whole sample	78,488	0.0093 (0.0260)	0.0176 (0.0271)	0.0075 (0.0297)			
Sample of wage earners	30,225	0.012 (0.0474)	0.024 (0.0481)	0.079 (0.0555)	0.0031 (0.00798)	0.00399 (0.00809)	0.0144 (0.00915)
Control variables:							
Year of birth*enrollment rate in 1971		No	Yes	Yes	No	Yes	Yes
Year of birth*water and sanitation program		No	No	Yes	No	No	Yes

Notes: All specifications include region of birth dummies, year of birth dummies, and interactions between the year of birth dummies and the number of children in the region of birth (in 1971). The number of observations listed applies to the specification in columns (1) and (4). Standard errors are in parentheses.

Figure by MIT OpenCourseWare.

Effect of the Program on Education and Wages: Coefficients of the Interactions Between Dummies Indicating Age in 1974 and the Number of Schools Constructed per 1,000 Children in Region of Birth

Age in 1974	*Dependent variable: years of education*						*Dependent variable: Log(hourly wage)*		
	Whole sample			*Sample of wage earners*					
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
12	-0.035	-0.025	0.002	-0.040	-0.010	0.009	0.016	0.019	0.027
	(0.047)	(0.048)	(0.054)	(0.077)	(0.078)	(0.091)	(0.013)	(0.013)	(0.015)
11	0.011	0.025	0.018	0.008	0.014	-0.003	-0.014	-0.013	-0.009
	(0.046)	(0.047)	(0.051)	(0.073)	(0.074)	(0.083)	(0.012)	(0.013)	(0.014)
10	0.059	0.049	0.078	0.10	0.092	0.13	0.0036	0.0042	0.0059
	(0.047)	(0.049)	(0.054)	(0.075)	(0.076)	(0.090)	(0.013)	(0.013)	(0.015)
9	0.14	0.14	0.15	0.067	0.063	0.17	0.0095	0.010	0.018
	(0.039)	(0.041)	(0.044)	(0.065)	(0.066)	(0.077)	(0.011)	(0.011)	(0.013)
8	0.088	0.11	0.11	0.19	0.20	0.28	0.019	0.021	0.027
	(0.049)	(0.050)	(0.054)	(0.078)	(0.079)	(0.089)	(0.013)	(0.013)	(0.015)
7	0.12	0.14	0.16	0.11	0.13	0.16	-0.0095	-0.0049	0.0066
	(0.044)	(0.046)	(0.051)	(0.072)	(0.073)	(0.084)	(0.012)	(0.012)	(0.014)
6	0.14	0.17	0.26	0.23	0.23	0.32	0.011	0.013	0.018
	(0.042)	(0.044)	(0.049)	(0.070)	(0.070)	(0.084)	(0.012)	(0.012)	(0.014)
5	0.10	0.13	0.13	0.14	0.16	0.27	0.021	0.023	0.052
	(0.043)	(0.045)	(0.050)	(0.075)	(0.075)	(0.088)	(0.013)	(0.013)	(0.015)
4	0.11	0.12	0.18	0.19	0.19	0.29	0.019	0.020	0.038
	(0.039)	(0.041)	(0.046)	(0.069)	(0.069)	(0.082)	(0.012)	(0.012)	(0.014)
3	0.11	0.14	0.20	0.15	0.17	0.30	0.0079	0.013	0.027
	(0.044)	(0.046)	(0.053)	(0.079)	(0.080)	(0.097)	(0.013)	(0.014)	(0.016)
2	0.14	0.19	0.19	0.20	0.22	0.25	0.016	0.023	0.040
	(0.041)	(0.043)	(0.049)	(0.073)	(0.074)	(0.088)	(0.012)	(0.013)	(0.015)
Control variables:									
Year of birth*enrollment rate in 1971	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year of birth*water and sanitation program	No	No	Yes	No	No	Yes	No	No	Yes
F-statistic	4.03	5.18	6.15	2.70	2.74	4.38	1.13	1.29	2.05
R^2	0.19	0.19	0.17	0.14	0.14	0.13	0.14	0.15	0.13
Number of observations	152,989	152,495	143,107	60,633	60,466	55,144	60,633	60,466	55,144

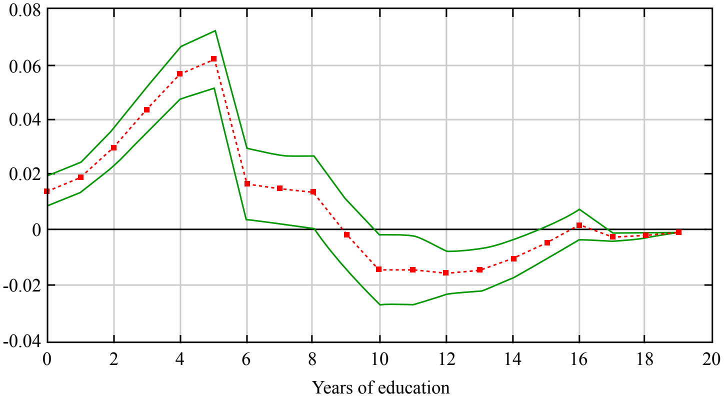
Figure by MIT OpenCourseWare.

Duflo (2001)

Effect of Education on Labor Market Outcomes: OLS and 2SLS Estimates

Method	*Instrument*	*(1)*	*(2)*	*(3)*
Panel A: Sample of wage earners				
Panel A1: Dependent variable: log(hourly wage)				
OLS		0.0776 (0.000620)	0.0777 (0.000621)	0.0767 (0.000646)
2SLS	Year of birth dummies*program intensity in region of birth	0.0675 (0.0280) [0.96]	0.0809 (0.0272) [0.9]	0.106 (0.0222) [0.93]
2SLS	(Aged 2-6 in 1974)*program intensity in region of birth	0.0752 (0.0338) (0.0338)	0.0862 (0.0336) (0.0336)	0.104 (0.0304) (0.0304)
Panel A2: Dependent variable: log(monthly earnings)				
OLS		0.0698 (0.000601)	0.0698 (0.000602)	0.0689 (0.000628)
2SLS	Year of birth dummies*program intensity in region of birth	0.0756 (0.0280) [0.73]	0.0925 (0.0278) [0.63]	0.0913 (0.0219) [0.58]
Panel B: Whole sample				
Panel B1: Dependent variable: participation in the wage sector				
OLS		0.0328 (0.00311)	0.0327 (0.000311)	0.0337 (0.000319)
2SLS	Year of birth dummies*program intensity in region of birth	0.101 (0.0210) [0.66]	0.118 (0.0197) [0.93]	0.0892 (0.0162) [1.12]
Panel B2: Dependent variable: log(monthly earnings), imputed for self-employed individuals				
OLS		0.0539 (0.000354)	0.0539 (0.000354)	0.0539 (0.000355)
2SLS	Year of birth dummies*program intensity in region of birth	0.0509 (0.0157) [0.68]	0.0745 (0.0136) [0.58]	0.0346 (0.0138) [0.16]

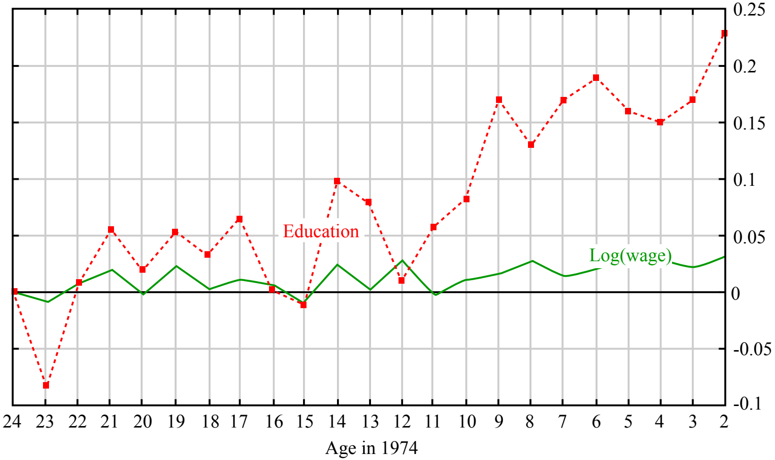
Figure by MIT OpenCourseWare.



Difference in differences in CDF (estimated from linear probability model) with 95-percent confidence interval

Figure by MIT OpenCourseWare.

Duflo (2001)



Coefficients of the interactions (age in 1974) * (program intensity in the region of birth) in the wage and education equations

Figure by MIT OpenCourseWare.

Duflo (2004)

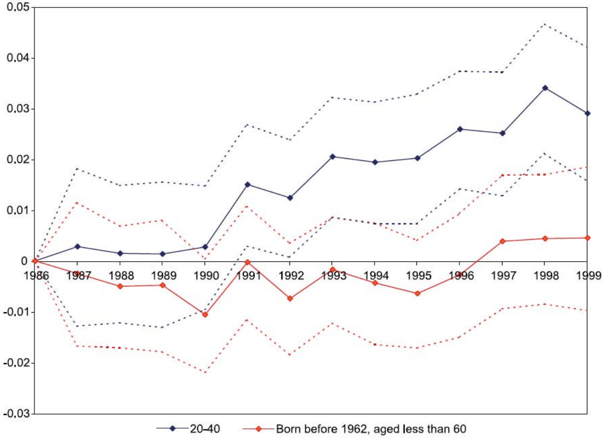


Fig. 2. Coefficients of the interactions of program intensity and survey year dummies. Dependent variable: % of primary school graduates.

Duflo (2004)

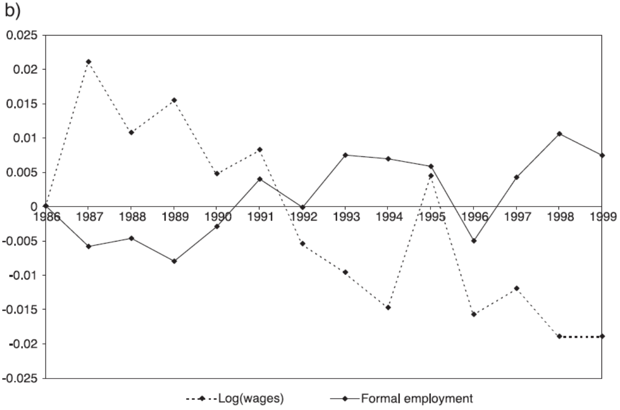


Fig. 4. (a) Coefficients of the interactions of program intensity and survey year dummies. Dependent variables: log(wage) and formal sector employment (individuals born before 1962 and aged less than 60). Sample: urban and rural regions. (b) Coefficients of the interactions of program intensity and survey year dummies. Dependent variables: average log(wage) and average formal sector employment among individuals born before 1962 and aged less than 60. Sample: rural regions.

Duflo (2004)

Table 6

2SLS estimates of the impact of average education on individual wages

	Independent variable: % of primary school graduates in the 20–40 sample		Independent variable: % of primary school graduates in the 20–60 sample	
	Sample: rural and urban areas	Sample: rural areas only	Sample: rural and urban areas	Sample: rural areas only
	(1)	(2)	(3)	(4)
Panel A: years 1986–1999				
Log (wage)	−0.204 (0.443)	−0.834 (0.701)	−0.208 (0.615)	−0.871 (0.837)
Log (wage) residual	−0.292 (0.355)	−0.633 (0.431)	−0.379 (0.512)	−0.994 (0.556)
Skill premium	−0.434 (0.916)	−0.982 (1.408)	−0.596 (1.197)	−0.636 (1.645)
Formal employment	0.441 (0.159)	0.454 (0.203)	0.661 (0.238)	0.745 (0.352)
Formal employment among educated workers	0.432 (0.197)	0.501 (0.259)	0.543 (0.264)	0.713 (0.406)
Formal employment among uneducated workers	0.379 (0.203)	0.409 (0.232)	0.510 (0.354)	0.318 (0.318)
Panel B: years 1986–1997				
Log (wage)	−0.358 (0.493)	−0.710 (0.821)	−0.451 (0.716)	−0.480 (0.801)
Log (wage) residual	−0.330 (0.412)	−0.588 (0.529)	−0.437 (0.618)	−0.902 (0.602)
Skill premium	−0.225 (1.033)	−0.635 (1.461)	−0.291 (1.488)	0.536 (1.576)
Formal employment	0.463 (0.183)	0.442 (0.233)	0.716 (0.282)	0.694 (0.379)
Formal employment among educated workers	0.428 (0.229)	0.473 (0.301)	0.530 (0.317)	0.622 (0.479)
Formal employment among uneducated workers	0.478 (0.249)	0.449 (0.277)	0.624 (0.415)	0.263 (0.319)

Men aged 20–60 and born before 1962.

1. Survey year dummies, region dummies, interactions between survey year dummies and the enrollment rate in 1971, and interactions between survey year dummies and the number of children are included in the regressions. 2. Regression run using kabupaten-year averages, weighted by the number of observations in each kabupaten-year cell. 3. The instruments are interactions between survey year dummies and the program intensity. 4. The standard errors are corrected for auto-correlation within kabupaten.

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

Education Quality

Esther Duflo

14.771

School quality in Developing Countries

- There has been rapid improvement in school enrollment in developing countries over the last 10-15 years.
- However these improvements have not been matched by improvement in school quality:
  - Low learning performance (ASER study in India)
  - Massive Teacher absence (Chaudhury and other: 24% in India)
- Education quality has been an extremely active domain of research, and in particular there are a series of randomized evaluation paper on various issues:
  - “Production function” issues: class size, textbooks, flipcharts, etc.
  - Incentives for students, parents, and teachers
  - School systems:
    - Pedagogy (curriculum etc.)
    - Para-teachers vs regular teachers
    - Parent information/mobilization (report cards, school commitees etc.)

Duflo, Hanna, Ryan: Incentives for Para-teachers

- In India, regular teachers have essentially no incentives (tenure, no increase in salary)
- Para-teachers and incentives
    - It should be easier to provide them with good incentives
    - However, in India, they are no more likely to be present
    - Could be because they are actually not provided with incentives
- Motivating questions for this paper:
    - Can an incentive programs for para-teachers increase their presence?
    - Would increase presence lead to increase in learning or would it be undermined by:
        - Multitasking
        - Loss in intrinsic motivation
        - Incompetence

What the paper does

1. A randomized Experiment in teacher incentives
2. A regression discontinuity Design scheme to interpret the results: We estimate the change in teacher behavior just before and just after the end of a month, and this suggests that they respond to financial incentives
3. Use the treatment group to estimate a structural model; The non-linear nature of the attendance rules allows for estimation of a simple dynamic labor supply model, where teacher chooses every day between going to school or staying home and getting an outside option

The Context

- We worked with Seva Mandir, an NGO in rural Rajasthan
- They run 150 "non-formal education center" (NFE): single teacher school for students who do not attend regular school.
- Students are 7-14 year old, completely illiterate when they join.
- Schools teach basic hindi and math skills and prepare students to "graduate" to primary school.
- In 1997, 20 million children were served by such NFEs

The Intervention

- Teacher in Intervention school were provided with a camera with non-temperable time and date stamp

A picture

Photograph of children in school removed due to copyright restrictions.

The Intervention

- Teacher in Intervention school were provided with a camera with non-temperable time and date stamp
- Instructed to take a picture of themselves and the children every day (morning and afternoon). A valid pairs of picture has:
  - Two pictures taken the same day, separated by at least 5 hours each.
  - At least 8 children per picture
- Payment is calculated each month and is a non-linear function of attendance:
  - Up to 10 days: Rs 500.
  - Each day above 10 days: Rs 50.
- In non-intervention schools, teachers receive Rs 1000, and are reminded by attending at least 20 days is compulsory.

The Evaluation

- We originally picked 120 schools, out of which 7 closed immediately after they were picked to be in the study (unrelated to the study).
- 57 treatment schools, the rest control.
- Data collection:
  - Teacher and student attendance: Monthly random checks.
  - In treatment schools: Camera data
  - Students learning: tests in September 03-April 04-Oct 04
  - Long term impact: a new sets of random checks was done in 2006-2007, and a new set of test scores were done in 2007

The Randomized evaluation Checklist

1. What was the power of the Experiment?
    - At what level was the experiment randomized?
    - We need to take into account clustering at that level in computing our standard error
    - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
    - Ways to enforce balance: Stratifying
    - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
    - If so, how did we deal with it?
4. Did we have non-compliance?
    - If so how did we deal with it?
5. Did we have contagion (externalities) between treatment and control group?

Power

- We know that $E[Y_i(0)|W_i = 1] = [Y_i(0)|W_i = 0]$
- But in a finite sample, it may or may not hold.
- Size (level) of a test (e.g. test H_0 ATE=0): Probability of a type I error: I reject H_0 when H_0 is true
- Generally we set the size at 5%.
- Power of a test: 1-probability of type II error.
- Type II error: for a given size, I do not reject 0, when I should have.
- Power depend on effect of program, and on precision of the estimate:
    - Sample size
    - Level of Randomization: If I randomize at the group level, I need to cluster at this group level: need to adjust power calculation for that (it will depend on size of the group, and expected correlation of outcomes within the group).

The Randomized evaluation Checklist

1. What was the power of the Experiment?
   - At what level was the experiment randomized?
   - We need to take into account clustering at that level in computing our standard error
   - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
   - Ways to enforce balance: Stratifying (creating block of covariates, and randomize within those)
   - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
   - If so, how did we deal with it?
4. Did we have non-compliance?
   - If so how did we deal with it?
5. Did we have contagion (externalities) between treatment and control group?

Checking the Balance in the Camera Experiment

Table 1: Is School Quality Similar in Treatment and Control Groups Prior to Program?

	Treatment (1)	Control (2)	Difference (3)
A. Teacher Attendance			
School Open	0.66	0.64	0.02
			(0.11)
	41	39	80
B. Student Participation (Random Check)			
Number of Students Present	17.71	15.92	1.78
			(2.31)
	27	25	52
C. Teacher Qualifications			
Teacher Test Scores	34.99	33.62	1.37
			(2.01)
	53	56	109
Teacher Highest Grade Completed	10.21	9.80	0.41
			(0.46)
	57	54	111

School quality

Table 1: Is School Quality Similar in Treatment and Control Groups Prior to Program?

	Treatment (1)	Control (2)	Difference (3)
D. Teacher Performance Measures (Random Check)			
Percentage of Children Sitting Within Classroom	0.83	0.84	0.00
			(0.09)
	27	25	52
Percent of Teachers Interacting with Students	0.78	0.72	0.06
			(0.12)
	27	25	52
Blackboards Utilized	0.85	0.89	-0.04
			(0.11)
	20	19	39
E. School Infrastructure			
Infrastructure Index	3.39	3.20	0.19
			(0.30)
	57	55	112
Fstat(1,110)			1.21
p-value			(0.27)

Students

Table 2: Are Students Similar Prior To Program?

	Levels			Normalized by Control		
	Treatment	Control	Difference	Treatment	Control	Differenc
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Can the Child Write?					
Took Written Exam	0.17	0.19	-0.02			
			(0.04)			
	1136	1094	2230			
	B. Oral Exam					
Math Score on Oral Exam	7.82	8.12	-0.30	-0.10	0.00	-0.10
			(0.27)			(0.09)
	940	888	1828	940	888	1828
Language Score on Oral Exam	3.63	3.74	-0.10	-0.03	0.00	-0.03
			(0.30)			(0.08)
	940	888	1828	940	888	1828
Total Score on Oral Exam	11.44	11.95	-0.51	-0.08	0.00	-0.08
			(0.48)			(0.07)
	940	888	1828	940	888	1828
	C. Written Exam					
Math Score on Written Exam	8.62	7.98	0.64	0.23	0.00	0.23
			(0.51)			(0.18)
	196	206	402	196	206	402
Language Score on Written Exam	3.62	3.44	0.18	0.08	0.00	0.08
			(0.46)			(0.20)
	196	206	402	196	206	402
Total Score on Written Exam	12.17	11.41	0.76	0.16	0.00	0.16
			(0.90)			(0.19)
	196	206	402	196	206	402

The Randomized evaluation Checklist

1. What was the power of the Experiment?
   - At what level was the experiment randomized?
   - We need to take into account clustering at that level in computing our standard error
   - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
   - Ways to enforce balance: Stratifying (creating block of covariates, and randomize within those)
   - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
   - If so, how did we deal with it?
4. Did we have non-compliance?
   - If so how did we deal with it?
5. Did we have contagion (externalities) between treatment and control group?

Attrition

- At the school level: some schools got lost, for reasons not related to the program
- At the individual level for the test: we have substantial attrition
  - Why is that a potential problem?
  - When will it be a problem?
  - What should we check?
    - percentage attrition is not differential by group
    - observable characteristics of attritors are no different in T and C group
  - If not what can we do?
    - Assume a selection process, and correct for it (we lose main advantage of a random sample)
    - Provide bounds

Attrition

Table 9: Descriptive Statistics for Mid Test and Post Test

	Mid Test			Post Test		
	Treatment	Control	Difference	Treatment	Control	Difference
	A. Attrition Process					
Percent Attrition	0.11	0.22	-0.10	0.24	0.21	0.03
			(0.05)			(0.04)
Difference in Percent Written of Pre-Test attriters-stayers	0.01	0.03	0.02	0.06	-0.03	0.10
			(0.06)			(0.06)
Difference in Verbal Test of Pre-Test attriters-stayers	0.05	0.08	-0.03	0.02	0.12	-0.10
			(0.14)			(0.14)
Difference in Written Test of Pre-Test attriters-stayers	-0.41	-0.23	-0.18	-0.19	-0.13	-0.06
			(0.34)			(0.29)
	B. Exam Score Means					
Took Written	0.36	0.33	0.03	0.61	0.57	0.04
			(0.04)			(0.05)
Math	0.14	0.00	0.14	-0.08	-0.24	0.16
			(0.10)			(0.15)
Language	0.14	0.00	0.14	1.71	1.60	0.11
			(0.10)			(0.11)
Total	0.14	0.00	0.14	0.35	0.24	0.12
			(0.10)			(0.11)

The Randomized evaluation Checklist

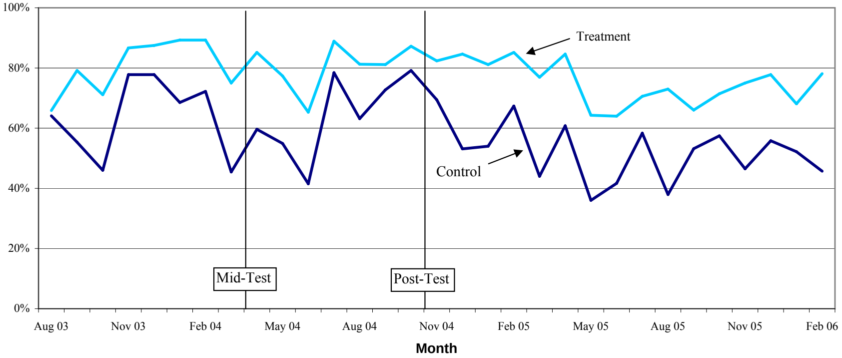
1. What was the power of the Experiment?
   - At what level was the experiment randomized?
   - We need to take into account clustering at that level in computing our standard error
   - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
   - Ways to enforce balance: Stratifying (creating block of covariates, and randomize within those)
   - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
   - If so, how did we deal with it?
4. Did we have non-compliance?
   - If so how did we deal with it? (next lecture)
5. Did we have contagion (externalities) between treatment and control group?

The Randomized evaluation Checklist

1. What was the power of the Experiment?
    - At what level was the experiment randomized?
    - We need to take into account clustering at that level in computing our standard error
    - This affect our *power* as well
2. What the randomization successful (was there balance between treatment and control group in covariates)
    - Ways to enforce balance: Stratifying (creating block of covariates, and randomize within those)
    - Ways to check balance: Compare covariates
3. Did we have attrition (lost observations)?
    - If so, how did we deal with it?
4. Did we have non-compliance?
    - If so how did we deal with it?
5. Did we have contagion (externalities) between treatment and control group?

Attendance: Graphical Evidence

Figure 2: Percentage of Schools Open during Random Checks



Attendance: tables

Table 3: Teacher Attendance

Sept 2003-Feb 2006			Difference Between Treatment and Control Schools		
Treatment (1)	Control (2)	Diff (3)	Until Mid-Test (4)	Mid to Post Test (5)	After Post Test (6)
			A. All Teachers		
0.79	0.58	0.21	0.20	0.20	0.23
		(0.03)	(0.04)	(0.04)	(0.04)
1575	1496	3071	882	660	1529
		B. Teachers with Above Median Test Scores			
0.78	0.63	0.15	0.15	0.15	0.14
		(0.04)	(0.05)	(0.05)	(0.06)
843	702	1545	423	327	795
		C. Teachers with Below Median Test Scores			
0.78	0.53	0.24	0.21	0.14	0.32
		(0.04)	(0.05)	(0.06)	(0.06)
625	757	1382	412	300	670

Cheating?

Table 4: Comparing Random Checks to Photo Data for Treatment Schools

Scenario	Number	Percent of Total
A. Possible Scenarios		
School Open and Valid Photos	879	66%
School Open and Invalid Photos	179	13%
School Closed and Valid Photos	88	7%
School Closed and Invalid Photos	191	14%
B. Out of 179 where School is Open, the photos are invalid because….		
School not open for full 5 hours	43	24%
Only one photo	90	50%
Not enough Children	36	20%
Instructor not in Photo	9	5%
Don't Know	1	1%
C. Out of 88 where School is Closed and the photos are valid…..		
Random check completed after the school closed	13	15%
Camera broke/excused meeting	21	24%
Teacher left in the middle of the day	54	61%

No evidence of Multitasking

Table 7: Teacher Performance

	Sept 2003-Feb 2006			Difference Between Treatment and Control Schools		
	Treatment	Control	Diff	Until Mid-Test	Mid to Post Test	After Post Test
	(1)	(2)	(3)	(4)	(5)	(6)
Percent of Children Sitting Within Classroom	0.72	0.73	-0.01	0.01	0.04	-0.01
			(0.01)	(0.89)	(0.03)	(0.02)
	1239	867	2106	643	480	983
Percent of Teachers Interacting with Students	0.55	0.57	-0.02	-0.02	0.05	-0.04
			(0.02)	(0.04)	(0.05)	(0.03)
	1239	867	2106	643	480	983
Blackboards Utilized	0.92	0.93	-0.01	-256766.00	0.01	-0.01
			(0.01)	(0.02)	(0.02)	(0.02)
	990	708	1698	613	472	613

Notes: (1) Teacher Performance Measures from Random Checks only includes schools that were open during the random check. (2) Standard errors are clustered by school.

No increase on conditional attendance, more days worked

Table 8: Child Attendance

	Sept 03-Feb 06			Difference Between Treatment and Control Schools		
	Treatment	Control	Diff	Until Mid-Test	Mid to Post Test	After Post Test
	(1)	(2)	(3)	(4)	(5)	(6)
A. Attendance Conditional on School Open						
Attendance of Students Present at Pre-Test Exam	0.46	0.46	0.01	0.02	0.03	0.00
			(0.03)	(0.03)	(0.04)	(0.03)
	23495	16280	39775			
Attendance for Children who did not leave NFE	0.62	0.58	0.04	0.02	0.04	0.05
			(0.03)	(0.03)	(0.04)	(0.03)
	12956	10737	23693			
B. Total Instruction Time (Presence)						
Presence for Students Present at Pre-Test Exam	0.37	0.28	0.09	0.10	0.10	0.08
			(0.03)	(0.03)	(0.04)	(0.03)
	29489	26695	56184			
Presence for Student who did not leave NFE	0.50	0.36	0.13	0.10	0.13	0.15
			(0.03)	(0.04)	(0.05)	(0.04)
	16274	17247	33521			
C. Presence, by Student Learning Level at Program Start (for those who did not leave)						
Took Oral Pre-Test	0.50	0.36	0.14	0.11	0.14	0.15
			(0.03)	(0.03)	(0.05)	(0.04)
	14778	14335	29113			
Took Written Pre-Test	0.48	0.39	0.10	0.07	0.07	0.11
			(0.06)	(0.07)	(0.06)	(0.07)
	1496	2912	4408			

Notes: (1) Standard errors are clustered at the level of the school. (2) Child attendance data were collected during random checks. (3) The attendance at the pre-test exam determined the child enrollment at the start of the program.

Regression

$$\text{Score}_{ikj} = \beta_1 + \beta_2 \text{Treat}_j + \beta_3 \text{Pre\_Writ}_{ij} + \beta_4 \text{Pre\_oral}_{ij} + \beta_5 \text{Writ} + \epsilon_{ijk}$$

Test Score results

Table 10: Estimation of Treatment Effects for the Mid- and Post-Test

Mid-Test				Post-Test			
Took Written (1)	Math (2)	Lang (3)	Total (4)	Took Written (5)	Math (6)	Lang (7)	Total (8)
A. All Children							
0.04	0.15	0.16	0.17	0.06	0.21	0.16	0.17
(0.03)	(0.07)	(0.06)	(0.06)	(0.04)	(0.12)	(0.08)	(0.09)
1893	1893	1893	1893	1760	1760	1760	1760
B. With Controls							
0.02	0.13	0.13	0.14	0.05	0.17	0.13	0.15
(0.03)	(0.07)	(0.05)	(0.06)	(0.04)	(0.10)	(0.07)	(0.07)
1893	1893	1893	1893	1760	1760	1760	1760

Results by Pre-test score

Table 10: Estimation of Treatment Effects for the Mid- and Post-Test

Mid-Test			**Post-Test**			
Math (2)	Lang (3)	Total (4)	Took Written (5)	Math (6)	Lang (7)	Total (8)
			C. Took Pre-Test Oral			
0.14	0.13	0.15		0.2	0.13	0.16
(0.08)	(0.06)	(0.07)		(0.14)	(0.09)	(0.10)
1550	1550	1550		1454	1454	1454
			D. Took Pre-Test Written			
0.19	0.28	0.25		0.28	0.28	0.25
(0.12)	(0.11)	(0.11)		(0.18)	(0.11)	(0.12)
343	343	343		306	306	306

Graduation to government school

Table 11: Dropouts and Movement into Government Schools

	Treatment (1)	Control (2)	Diff (3)
Child Left NFE	0.44	0.36	0.08
			(0.04)
Child Enrolled in Government School	0.26	0.16	0.10
			(0.03)
Child Dropped Out of School	0.18	0.20	-0.02
			(0.03)
N	1136	1061	2197

Estimating the impact of teacher absence

- Suppose we want to use this experiment to estimate the impact of teacher absence on test score?
- What would the strategy be?
  - Use Treatment dummy as instrument for teacher attendance
  - Wald estimate: divide effect of program on test score by effect of program on attendance
- What would the potential threat to validity of the strategy
- What do we think about the severity of this threat?

Estimating the impact of teacher absence

Table 12: Does the Random Check Predict Test Scores?

Method:	OLS	OLS	OLS	2SLS
Sample:	Control Schools	Treatment Schools	Treatment Schools	All Schools
Data:	Random Check	Random Check	Photographs	Random Check
	(1)	(2)	(3)	(4)
		A. Mid-test (Sept 03-April 04)		
Took Written	0.02	0.28	0.36	0.26
	(0.10)	(0.08)	(0.11)	(0.19)
Total Score	0.20	0.39	0.87	1.07
	(0.19)	(0.21)	(0.22)	(0.43)
N	878	1015	1015	1893
		B. Post-test (Sept 03 -Oct 04)		
Took Written	0.24	0.51	0.59	0.33
	(0.16)	(0.15)	(0.20)	(0.22)
Total Score	0.58	1.17	0.98	0.97
	(0.35)	(0.36)	(0.53)	(0.47)
N	883	877	877	1760

Monitoring or Incentives? Preliminary Evidence

- Are teachers sensitive to increased monitoring or to incentives?
- Preliminary evidence based on *Regression Discontinuity Design*
- Consider a case where treatment is assigned when the treatment is assigned based on a strict threshold:
  - Sharp RD: $W_i = 1[X_i > c]$
  - Fuzzy RD: $lim_{x \downarrow c} pr(W_i = 1 | Xi = x) \neq lim_{x \uparrow c} pr(W_i = 1 | Xi = x)$
- Identification assumption for RD: $lim_{x \downarrow c} E[Y_i(0) | Xi = x] = lim_{x \uparrow c} E[Y_i(0) | Xi = x]$
- Estimator: we try to approximate: $lim_{x \downarrow c} E[Y_i | Xi = x] - lim_{x \uparrow c} E[Y_i | Xi = x]$
  - In the sharp RD: this will be the treatment effect
  - In the fuzzy RD: we use the treshold as instrument: compute our friend the Wald estimate.

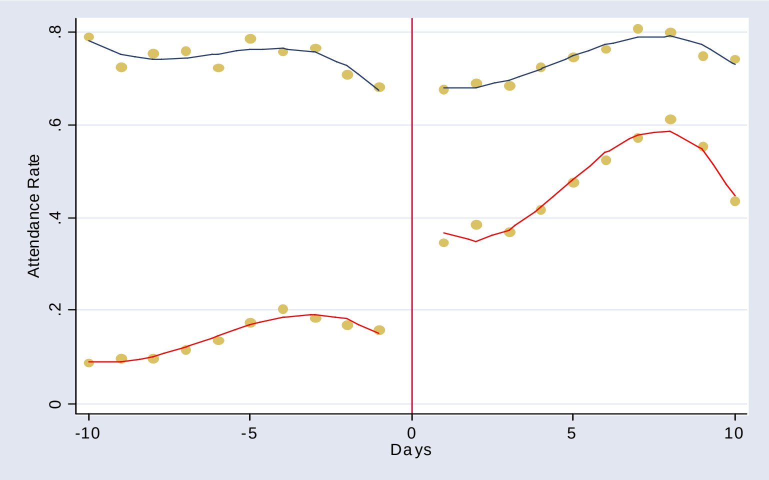
RD in the teacher case

- In practice: We try to estimate a smooth (non-parametric) function of the relationship between Y and X (here: day in the month and whether teacher works).
- We then use this to estimate the limits at the threshold, from the left and the right.
- When we switch from the last day of the month to the first day of the month:
    - A teacher who has attended 9 days or less in the rest of the month faces no incentive at the end of month t and faces incentives again at the end of month $t+1$.
    - A teacher who has attended more than 10 days in the rest of the month face a Rs 50 incentives at the end of month t and slightly smaller at the beginning of the next month
- Graphical Evidence
- Regression:

$$W_{itm} = \alpha + \beta 1_m(d > 10) + \gamma F + \lambda 1_m(d > 10) * F + \upsilon_i + \mu_m \epsilon_{is}, \tag{1}$$

Regression Discontinuity Design: Graphical Evidence

Figure 5: RDD Representation of Teacher Attendance at the Start and End of the Month



Regression Discontinuity Design: Regressions

Table 5 : Do Teachers Work More When They are "In the Money"?

	(1)	(2)	(3)	(4)
Beginning of Month	0.19	0.12	0.46	0.39
	(0.05)	(0.06)	(0.04)	(0.03)
In the Money	0.52	0.37	0.6	0.48
	(0.04)	(0.05)	(0.03)	(0.01)
Beginning of the Month * In the Money	-0.19	-0.12	-0.34	-0.3
	(0.06)	(0.06)	(0.04)	(0.02)
Observations	2813	2813	27501	27501
R-squared	0.06	0.22	0.08	0.16
Sample	1st and last day of month	1st and last day of month	1st 10 and last 10 days of month	1st 10 and last 10 days of month
Third Order Polynomial on Days on each side			X	X
Teacher Fixed Effects		X		X
Month Fixed Effects		X		X
Clustered Standard Errors	X		X	

The Model

- Each day, a teacher chooses whether or not to attend school, by comparing the value of attending school to that of staying home or doing something else.
- State space $s = (t, d)$, where t is the current time and d is the days worked previously in the current month.
- Payoffs:
  - If the teacher does not attend school: $\mu + \epsilon_t$
  - Payoff of attending school is calculated at the end of the month according to:

$$\pi(d) = 500 + \max\{0, d - 10\} \tag{2}$$

- T takes value between 1 and $T = 25$.
- Transitions: Each day, t increases by one, unless $t = T$, in which case it resets to $t = 1$. If a teacher has worked in that period d increases by one, otherwise it remains constant.

Value function

Given this payoff structure, for $t < T$, we can write the value function for each teacher as follows:

$$V(t, d) = \max\{\mu + \epsilon_t + EV(t+1, d), EV(t+1, d+1)\}. \quad (3)$$

At time T, we have:

$$V(T, d) = \max\{\mu + \epsilon_T + \beta\pi(d) + EV(1, 0), \beta\pi(d+1) + EV(1, 0)\}, \quad (4)$$

where β is marginal utility of income.
$EV(1, 0)$ enters both side and can thus be ignored: we can solve each month independently, backwards from time T.

Identification

- Identification is constructive, and based on partitions of the state space.
- At time T, the agent faces a static decision; work if:

$$\mu + \epsilon_T + \beta\pi(d) > \beta\pi(d+1). \tag{5}$$

- The probability of this event is:

$$\begin{aligned} Pr(work|d,\theta) &= Pr(\epsilon_T > \beta(\pi(d+1) - \pi(d)) - \mu) \quad (6) \\ &= 1.0 - \Phi(\beta(\pi(d+1) - \pi(d)) - \mu), \quad (7) \end{aligned}$$

Identification with iid innovation in outside option

- When $d < 10$, the difference between $\pi(d+1)$ and $\pi(d)$ is zero, and β does not enter the equation.
- The resulting equation is:

$$Pr(work|d, \theta) = 1 - \Phi(\mu), \tag{8}$$

which is a simple probit.

- If all teachers share same μ, μ is identified by teachers who are out of the money, and then β from teachers in the money.
- var(ϵ) normalized to be equal to 1.
- If teachers have different μ model still identified by comparing different teachers with themselves over time (teacher fixed effect).

Identification with AR(1) innovation in outside option

- If ϵ is serially correlated, identification is more complicated.
- Suppose that the shock follows an AR(1) process:

$$\epsilon_t = \rho\epsilon_{t-1} + \nu_t, \tag{9}$$

- ϵ_T will be correlated with d, as teachers with very high draws on ϵ_T are more likely to be in the region where $d < 10$ if ρ is positive (the converse will be true if ρ is negative).
- This will bias our estimates of μ and β.

iid model, with or without fixed effect

Simply write the empirical counterpart of the maximization problem.
The log likelihood is:

$$LLH(\theta) = \sum_{i=1}^{N} \sum_{m=1}^{M_i} \sum_{t=1}^{T_m} [1(\text{work}) Pr(\text{work}|t, d, \theta)$$

$$+1(\text{not work})(1 - Pr(\text{work}|t, d, \theta)],$$

where:

$$\begin{aligned} Pr(\text{work}|t, d, \theta) &= Pr(\mu + \epsilon_t + EV(t+1, d) < EV(t+1, d+1)) \\ &= Pr(\epsilon_t < EV(t+1, d+1) - EV(t+1, d) - \mu) \\ &= \Phi(EV(t+1, d+1) - EV(t+1, d) - \mu), \quad (10) \end{aligned}$$

Serial correlation

- Both estimation and identification are a little complicated...
- Use method of simulated moment: simulate work history for different parameters, and try to match a distribution of days worked at the beginning of the month.
- Can introduce heterogeneity by drawing p teacher from a distribution with high outside option, and $1 - p$ from distribution with low outside option.

Results from the structural Model

Table 6: Results from the Structural Model

Parameter	Model I (1)	Model II (2)	Model III (3)	Model IV (4)	Model V (5)	Model VI (6)
β	0.049	0.024	0.059	0.051	0.014	0.019
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
μ_1	1.55		2.315	2.063	-0.107	0.012
	(0.013)		(0.013)	(0.012)	(0.040)	(0.028)
ρ			0.682	0.547	0.461	
			(0.010)	(0.023)	(0.039)	
σ_1^2				0.001	0.153	0.135
				(0.011)	(0.053)	(0.027)
μ_2					3.616	1.165
					(0.194)	(0.101)
σ_2^2					0.26	0.311
					(0.045)	(0.051)
p					0.047	0.131
					(0.007)	(0.015)
Heterogeneity	None	FE	None	RC	RC	RC

Prediction on days worked (real=20.23 days)

Table 6: Results from the Structural Model

Parameter	Model I (1)	Model II (2)	Model III (3)	Model IV (4)	Model V (5)	Model VI (6)
Heterogeneity	None	FE	None	RC	RC	RC
ϵ_{Bonus}	3.52	1.687	6.225	10.08	0.306	0.370
	(1.550)	(0.098)	(0.634)	(1.249)	(0.038)	(0.029)
$\epsilon_{\text{bonus\_cutoff}}$	-75.49	-16.04	-50.22	-63.11	-1.29	-1.78
	(6.506)	(1.264)	(2.612)	(3.395)	(0.479)	(0.449)
Predicted Days Worked	20.50	19.00	15.30	12.15	20.23	21.36
	(0.031)	(0.062)	(0.058)	(0.102)	(3.512)	(0.373)
Days Worked BONUS=0	1.60	6.02	1.29	1.318	13.55	11.81
	(0.597)	(0.234)	(0.875)	(0.863)	(5.251)	(0.669)
Out of Sample Prediction	26.16	18.886	15.08	12.956	20.86	21.57
	(0.059)	(0.253)	(0.635)	(0.520)	(3.793)	(0.456)

Distribution of Days worked

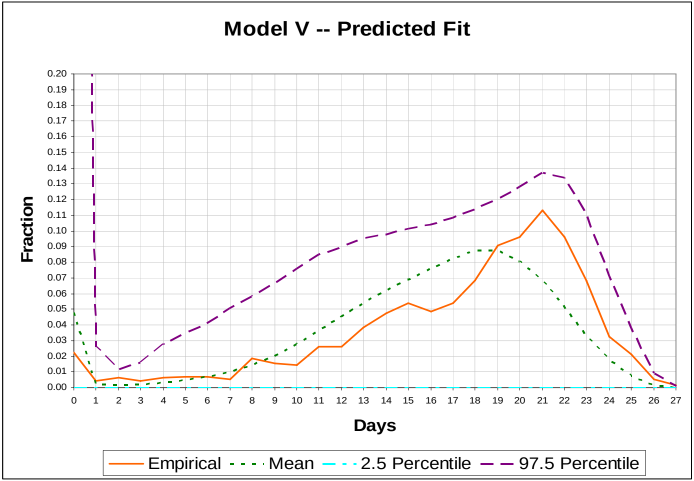


Figure 6B: CounterFactual Fit From Model V

Two out of sample tests

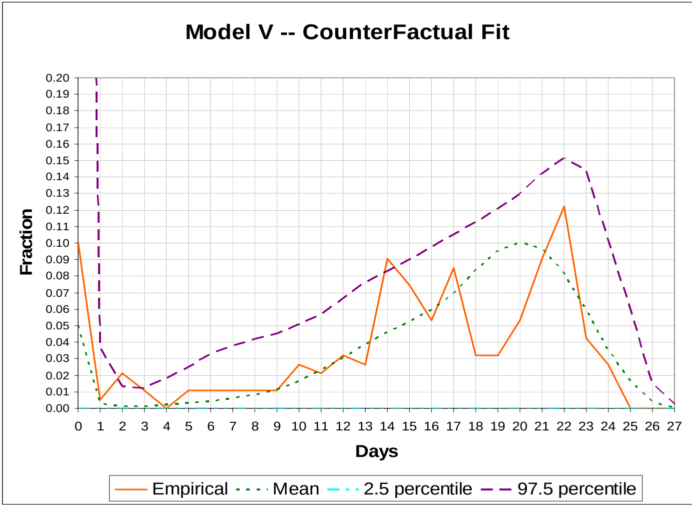
- Prediction of the number of days worked under no incentives
  - Model predicts that teachers would work 52% of the time in control group
  - In fact they work 58%
  - Predicted difference treatment vs control is 26%, vs 21% in reality
- The impact of a change in rule.
  - Seva Mandir changed rule after experiment was over (and model was estimated!)
  - New rule: Rs 700 for 12 days of work. Increment of Rs 70 after the 13th day
  - Model does well too.
- Note that all the alternative models do rather poorly in these counterfactuals.

Results from the structural Model

Table 6: Results from the Structural Model

Parameter	Model I (1)	Model II (2)	Model III (3)	Model IV (4)	Model V (5)	Model VI (6)
Heterogeneity	None	FE	None	RC	RC	RC
$\in_{\text{Bonus}}$	3.52	1.687	6.225	10.08	0.306	0.370
	(1.550)	(0.098)	(0.634)	(1.249)	(0.038)	(0.029)
$\in_{\text{bonus\_cutoff}}$	-75.49	-16.04	-50.22	-63.11	-1.29	-1.78
	(6.506)	(1.264)	(2.612)	(3.395)	(0.479)	(0.449)
Predicted Days Worked	20.50	19.00	15.30	12.15	20.23	21.36
	(0.031)	(0.062)	(0.058)	(0.102)	(3.512)	(0.373)
Days Worked BONUS=0	1.60	6.02	1.29	1.318	13.55	11.81
	(0.597)	(0.234)	(0.875)	(0.863)	(5.251)	(0.669)
Out of Sample Prediction	26.16	18.886	15.08	12.956	20.86	21.57
	(0.059)	(0.253)	(0.635)	(0.520)	(3.793)	(0.456)

Distribution of Days worked under new rule



Results from the structural model: Lessons

- A nice set up where we can corroborate assumptions of structural model.
- Other example: Todd and Wolpin (AER). They estimate a structural model in the control group and then validate it by predicting the Treatment Control difference.
- Model incorporating both serial correlation and heterogeneity does well, other models do poorly
- It seems that entire effect of program was through financial incentives.

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

Education Policy in Equilibrium

Esther Duflo

14.771

Four Examples were Looking at the Market Equilibrium Makes a Difference

- Private and Social Returns (done!)
- Vouchers for Private School
- Using Regression discontinuity to estimate the impact of class size
- The Cost of Teachers

Vouchers for private schools

- Why should government finance education?
- What should government provide education?
- Vouchers: A way to de-couple financing and provision of education
- The impact of vouchers: a partial equilibrium analysis Angrist et al (2006,2008)

Voucher in Partial Equilibrium

- Setting: one city in Colombia
- 125,000 Vouchers are randomly allocated to eligible students who applied through a lottery
- Vouchers cover half the cost of private schools
- They are renewed conditional on good performance in school
- Since Vouchers are assigned by lottery, we can follow losers and winners to asses their impact on probability to go to to private school, and their performance.
- First Stage and Reduced form impact of winning: Table 3
- Short term test score results of winning: Table 6
- Long term test impact: More likely to graduate high school and take school leaving exam Table 2

Dealing with Attrition and Non Compliance

- Attrition:
  - The differential rate at which students take the exam create an attrition problem
  - They deal with attrition in different ways, notably by constructing bounds
  - Without controlling: no difference in test scores. With Control: big differences on test scores
- The non compliance problem:
- Some kids do not take the voucher (never takers)
- Some kids get another voucher (always takers) use winning the lottery as an instrument for getting the private school voucher Table 8
  - Are the identification assumption satisfied?
  - Suppose that we were trying to use this instrument to measure the impact of attending private school, would they be satisfied?

Equilibrium Effects

- Why might the effect of the voucher on a winner not tell us what would the effect be of introducing vouchers in an entire school system?
- If we could do a giant randomized experiment on this, what would it be?
- Hsieh and Urquiola examine the case of Chile
  - In 1981, Chile introduced a nationwide voucher systems
  - Massive entry of private school : ▸ Graph
  - Especially in richer and more urban area
  - DD (panel) analysis of the impact of school markets that got more private schools: regress change in test scores on change on fraction of students in private school.
  - ▸ Results More Private school not associated with any increase in test score
  - Potential Confounding factors?

Sorting

- Explanation: market for schools.
  - Parents like good schools
  - They think good school=good test scores
  - They look for school with good peers (even if there are no peer effects)
  - Increased sorting Table5
  - Evidence that parents have strong willingness to pay for schools with good peers (even with low value added): Zhang (2008) find no effect of elite school on test score in China (lottery), even though parents are willing to pay a lot of money to attend these schools.

Estimating Class Size: Angrist and Lavy

- A classic RD design
- Israel: Class size should not be over 40
- This creates discontinuity in class size at multiple of 40:
- A Fuzzy RD, since the class size is not exactly following the rule.
- Therefore we use predicted class size (based on the rule) as an instrument for actual class size, controlling for smooth function of the class size.
- Results: Figure First Stage and Reduced forms 2SLS

What happens to this if Class Size is a choice:? Urquiola and Verhoogen

- Back to Chile...
- Schools are subject to a class size cap (45), and an integer constraints on the number of classroom: they respect this figure 5
- But they can chose how many kids they enroll, as well as the fee they set.
- Profit-maximizing schools will endogenous choose enrollment and fees to avoid being on the right side of a discontinuity: this will generate bunching figure 7
- Except if they are targeting parents who really value class size and are willing to pay more for it: they will then raise the fees.

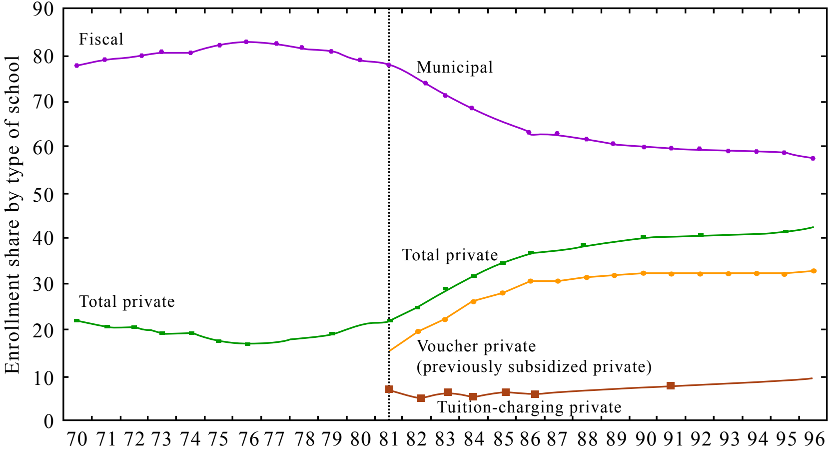
Problems for RD design

- The composition of students on the left and right side of a discontinuity will change endogenous as a function of the discontinuity ▸ figure 8
- This is a violation of the RD identification assumption: the potential outcome of the students will also differ.
- There is a RD in test scores ▸ figure 6 but it may be due to the underlying discontinuity in potential outcomes.
- ▸ Table 3: nicely estimated effect of class size
- ▸ Table 5: controlling for observable differences erase the result!
- Note that this does not invalidate the Angrist Lavy study, since in their set-up, class size was not a choice variables for the school.

The cost of Teachers

- Suppose returns to education increases (i.e. because of economic growth)
- This also will bid up the price of teachers
- And thus the cost of education
- This may not lead to an increase in education in steady state (Banerjee, 2004)
- Effect of education policy in equilibrium depends on:
  - What we assume about credit constraints
  - What we assume about preferences

Hsieh and Urquiola (2006)



National enrollment shares by sector, 1970-1996. Data assembled from several issues of the Ministry of Education's *Compendio Estadistico*.

Figure by MIT OpenCourseWare.

Table 3
OLS regressions for achievement, 1982–1988 and 1982–1996

	Dependent variable—change in average					
	Language score[a]			Math score[a]		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A—1982–1988						
Change in priv. enrollment[b]	−5.5	−6.7	−3.4	−7.2	−9.4	−9.2
	(7.5)	(7.7)	(8.7)	(7.6)	(7.5)	(8.9)
	[−0.08]	[−0.10]	[−0.05]	[−0.10]	[−0.13]	[−0.12]
Controls: previous trends[d]	No	Yes	Yes	No	Yes	Yes
Controls: concurrent trends[e]	No	No	Yes	No	No	Yes
N	84	84	84	84	84	84
R^2	0.006	0.073	0.105	0.010	0.087	0.156
Panel B—1982–1996						
Change in priv. enrollment[b]	−13.8*	−12.3	−8.9	−15.8**	−15.0**	−12.8
	(7.9)	(7.7)	(9.9)	(6.5)	(6.7)	(8.0)
	[−0.24]	[−0.21]	[−0.15]	[−0.27]	[−0.25]	[−0.22]
Controls: previous trends[d]	No	Yes	Yes	No	Yes	Yes
Controls: concurrent trends[e]	No	No	Yes	No	No	Yes
N	84	84	84	84	84	84
R^2	0.056	0.106	0.145	0.072	0.117	0.171

Table 5
Sorting among communes, 1990's cross-section and 1982–1988 changes

	Dependent variable—within commune observations of average characteristic in public schools/average characteristic in all schools									
	SES index[a]		Income[b]		Language[a]		Mathematics[a]		Repetition[c]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A—1990's cross sections[d]										
Private enrollment[c]	−0.20***	−0.16***	−0.37***	−0.33***	−0.08***	−0.08***	−0.09***	−0.09***	0.42***	0.28***
	(0.02)	(0.03)	(0.07)	(0.09)	(0.02)	(0.02)	(0.02)	(0.03)	(0.07)	(0.07)
	[−0.58]	[−0.46]	[−0.43]	[−0.38]	[−0.39]	[−0.39]	[−0.42]	[−0.42]	[0.44]	[0.29]
Commune controls[e]	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Thirteen regional dummies	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	296	296	184	184	296	296	296	296	299	299
R^2	0.313	0.493	0.171	0.285	0.188	0.396	0.215	0.346	0.193	0.447
Panel B—1982–1988 changes										
Change in private enrollment[c]					−0.21**	−0.22**	−0.14*	−0.19**	0.51**	0.38*
					(0.10)	(0.10)	(0.08)	(0.08)	(0.24)	(0.24)
					[−0.24]	[−0.26]	[−0.17]	[−0.23]	[0.24]	[0.18]
Controls: concurrent trends[f]					No	Yes	No	Yes	No	Yes
N					84	84	84	84	163	163
R^2					0.060	0.065	0.027	0.097	0.054	0.100

Angrist et al (2001)

Educational Outcomes and Voucher Status

Dependent variable	Bogota 95				Combined sample	
	Loser's means	No Ctls	Basic Ctls	Basic +19 Barrio Ctls	Basic Ctls	Basic +19 Barrio Ctls
	(1)	(2)	(3)	(4)	(5)	(6)
Using any scholarship in survey year	.057 (.232)	.509** (.023)	.504** (.023)	.505** (.023)	.526** (.019)	.521** (.019)
Ever used a scholarship	.243 (.430)	.672** (.021)	.663** (.022)	.662** (.022)	.636** (.019)	.635** (.019)
Started 6^{th} in private	.877 (.328)	.063** (.017)	.057** (.017)	.058** (.017)	.066** (.016)	.067** (.016)
Started 7^{th} in private	.673 (.470)	.174** (.025)	.168** (.025)	.171** (.024)	.170** (.021)	.173** (.021)
Currently in private school	.539 (.499)	.160** (.028)	.153** (.027)	.156** (.027)	.152** (.023)	.154** (.023)
Highest grade completed	7.5 (.960)	.164** (.053)	.130** (.051)	.120** (.051)	.085** (.041)	.078* (.041)
Currently in school	.831 (.375)	.019 (.022)	.007 (.020)	.007 (.020)	-.002 (.016)	-.002 (.016)
Finished 6^{th} grade	.943 (.232)	.026** (.012)	.023* (.012)	.021* (.011)	.014 (.011)	.012 (.010)
Finished 7^{th} grade (excludes Bog 97)	.847 (.360)	.040** (.020)	.031 (.019)	.029 (.019)	.027 (.018)	.025 (.018)
Finished 8^{th} grade (excludes Bog 97)	.632 (.483)	.112** (.027)	.100** (.027)	.094** (.027)	.077** (.024)	.074** (.024)
Repetitions of 6^{th} grade	.194 (.454)	-.066** (.024)	-.059** (.024)	-.059** (.024)	-.049** (.019)	-.049** (.019)
Ever repeated after lottery	.224 (.417)	-.060** (.023)	-.055** (.023)	-.051** (.023)	-.055** (.019)	-.053** (.019)
Total repetitions since lottery	.254 (.508)	-.073** (.028)	-.067** (.027)	-.064** (.027)	-.058** (.022)	-.057** (.022)
Years in school since lottery	3.7 (.951)	.058 (.052)	.034 (.050)	.031 (.050)	.015 (.044)	.012 (.043)
Sample size	562		1147		1577	

Figure by MIT OpenCourseWare.

Angrist et al (2001)

Test Results

	OLS estimates (1)	OLS with covariates (2)	RE estimates (3)	RE with covariates (4)
A. Test scores for all applicants				
Total points	.217* (.120)	.205* (.108)		
Math scores	.178 (.118)	.153 (.110)		
Reading scores	.204* (.122)	.203* (.113)		
Writing scores	.126 (.121)	.128 (.114)		
Pooled test scores			.170* (.095)	.148* (.088)
Math and reading scores			.192* (.101)	.162* (.096)
B. Test scores for female applicants				
Total points	.199 (.152)	.263** (.120)		
Math scores	.292** (.142)	.346** (.126)		
Reading scores	.117 (.156)	.152 (.135)		
Math and reading scores			.204 (.130)	.235** (.117)
B. Test scores for male applicants				
Total points	.204 (.183)	.170 (.179)		
Math scores	.010 (.142)	.004 (.031)		
Reading scores	.276 (.190)	.220 (.176)		
Math and reading scores			.143 (.160)	.087 (.160)

Figure by MIT OpenCourseWare.

Angrist et al (2001)

OLS and 2SLS Estimates of the Effect of Ever Using a Private School Scholarship

Dependent variable	Coefficient on ever used a private school scholarship				
	Bogota 95			Combined sample	
	Loser's mean (1)	OLS (2)	2SLS (3)	OLS (4)	2SLS (5)
Highest grade completed	7.5 (.965)	.167** (.053)	.196** (.078)	.141** (.042)	.134** (.065)
In school	.831 (.375)	.021 (.021)	.010 (.031)	.033* (.017)	-.003 (.026)
Total repetitions since lottery	.254 (.508)	-.077** (.029)	-.100** (.042)	-.069** (.023)	-.091** (.035)
Finished 8th	.632 (.483)	.114** (.028)	.151** (.041)	.108** (.025)	.127** (.038)
Test scores - total points	-.099 (1.0)	.379** (.111)	.291* (.153)	--	--
Married or living w/companion	.016 (.126)	-.009 (.006)	-.013 (.009)	-.010* (.006)	-.014 (.009)
N	562	1147		1577	

Figure by MIT OpenCourseWare.

Table 2. Voucher Status and the Probability of ICFES Match

	Exact ID Match		ID and City Match		ID and 7-letter Name Match		ID, City, and 7-letter Match	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. All Applicants (N=3542)							
Dependent Var. Mean	.354		.339		.331		.318	
Voucher Winner	.072	.059	.069	.056	.072	.059	.068	.056
	(.016)	(.015)	(.016)	(.014)	(.016)	(.014)	(.016)	(.014)
Male		-.052		-.053		-.043		-.045
		(.014)		(.014)		(.014)		(.014)
Age		-.160		-.156		-.153		-.149
		(.005)		(.005)		(.005)		(.005)
	B. Female Applicants (N=1789)							
Dependent Var. Mean	.387		.372		.361		.348	
Voucher Winner	.067	.056	.069	.057	.071	.060	.073	.062
	(.023)	(.021)	(.023)	(.021)	(.023)	(.021)	(.023)	(.021)
Age		-.168		-.164		-.160		-.156
		(.006)		(.006)		(.006)		(.006)
	C. Male Applicants (N=1752)							
Dependent Var. Mean	.320		.304		.302		.288	
Voucher Winner	.079	.063	.071	.055	.074	.059	.065	.050
	(.022)	(.020)	(.022)	(.020)	(.022)	(.020)	(.022)	(.020)
Age		-.153		-.148		-.146		-.141
		(.007)		(.007)		(.007)		(.006)

Notes. Robust standard errors are shown in parentheses. The sample includes all Bogotá 95 applicants with valid ID numbers and valid age data (i.e. ages 9 to 25 at application). The sample is the same as in Table 1, Column 5.

Angrist and Lavy (1999)

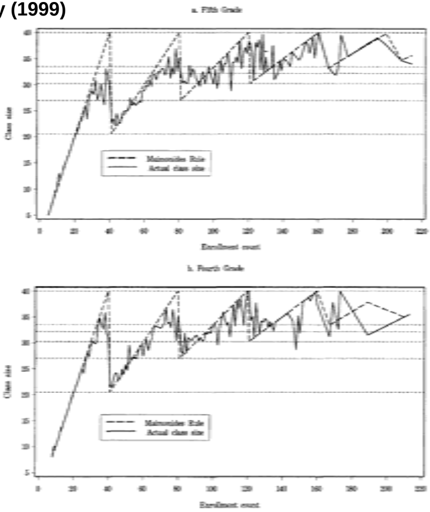


FIGURE I
Class Size in 1991 by Initial Enrollment Count, Actual Average Size and as Predicted by Maimonides' Rule

Angrist and Lavy (1999)

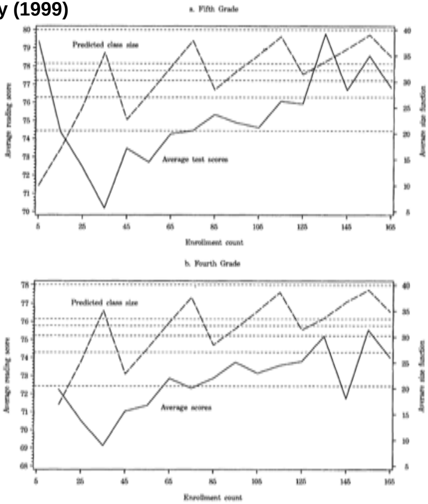


FIGURE II
Average Reading Scores by Enrollment Count, and the Corresponding Average Class Size Predicted by Maimonides' Rule

Angrist and Lavy (1999)

TABLE III
REDUCED-FORM ESTIMATES FOR 1991

	5th Graders						4th Graders					
	Class size		Reading comprehension		Math		Class size		Reading comprehension		Math	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Full sample												
Means	29.9		74.4		67.3		30.3		72.5		68.9	
(s.d.)	(6.5)		(7.7)		(9.6)		(6.3)		(8.0)		(8.8)	
Regressors												
f_{sc}	.704	.542	−.111	−.149	−.009	−.124	.772	.670	−.085	−.089	.038	−.033
	(.022)	(.027)	(.028)	(.035)	(.039)	(.049)	(.020)	(.025)	(.031)	(.040)	(.037)	(.047)
Percent disadvantaged	−.076	−.053	−.360	−.355	−.354	−.338	−.054	−.039	−.340	−.340	−.292	−.282
	(.010)	(.009)	(.012)	(.013)	(.017)	(.018)	(.008)	(.009)	(.013)	(.014)	(.016)	(.016)
Enrollment		.043		.010		.031		.027		.001		.019
		(.005)		(.006)		(.009)		(.005)		(.007)		(.009)
Root MSE	4.56	4.38	6.07	6.07	8.33	8.28	4.20	4.13	6.64	6.64	7.83	7.81
R^2	.516	.553	.375	.377	.247	.255	.561	.575	.311	.311	.204	.207
N	2,019		2,019		2,018		2,049		2,049		2,049	
B. Discontinuity sample												
Means	30.8		74.5		67.0		31.1		72.5		68.7	
(s.d.)	(7.4)		(8.2)		(10.2)		(7.2)		(7.8)		(9.1)	
Regressors												
f_{sc}	.481	.346	−.197	−.202	−.089	−.154	.625	.503	−.061	−.075	.059	.012
	(.053)	(.052)	(.050)	(.054)	(.071)	(.077)	(.050)	(.053)	(.056)	(.063)	(.072)	(.080)
Percent disadvantaged	−.130	−.067	−.424	−.422	−.435	−.405	−.068	−.029	−.348	−.343	−.306	−.291
	(.029)	(.028)	(.027)	(.029)	(.039)	(.042)	(.029)	(.028)	(.032)	(.034)	(.041)	(.043)
Enrollment		.086		.003		.041		.063		.007		.024
		(.015)		(.015)		(.022)		(.014)		(.017)		(.022)
Root MSE	5.95	5.58	6.24	6.24	8.58	8.53	5.49	5.26	6.57	6.57	8.26	8.25
R^2	.360	.437	.421	.421	.296	.305	.428	.475	.299	.299	.178	.182
N	471		471		471		415		415		415	

The function f_{sc} is equal to enrollment/[int((enrollment − 1)/40) + 1]. Standard errors are reported in parentheses. Standard errors were corrected for within-school correlation between classes. The unit of observation is the average score in the class.

Angrist and Lavy (1999)

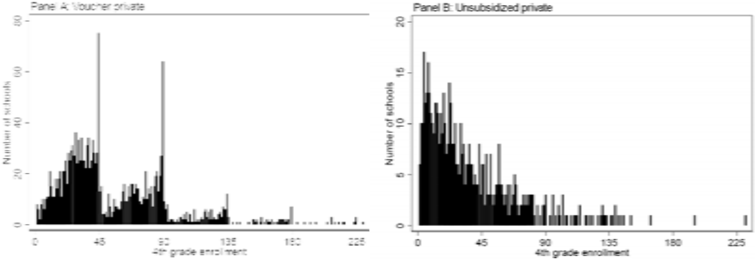
TABLE IV
2SLS Estimates for 1991 (Fifth Graders)

	Reading comprehension						Math					
	Full sample				+/− 5 Discontinuity sample		Full sample				+/− 5 Discontinuity sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mean score	74.4				74.5		67.3				67.0	
(s.d.)	(7.7)				(8.2)		(9.6)				(10.2)	
Regressors												
Class size	−.158	−.275	−.260	−.186	−.410	−.582	−.013	−.230	−.261	−.202	−.185	−.443
	(.040)	(.066)	(.081)	(.104)	(.113)	(.181)	(.056)	(.092)	(.113)	(.131)	(.151)	(.236)
Percent disadvantaged	−.372	−.369	−.369		−.477	−.461	−.355	−.350	−.350		−.459	−.435
	(.014)	(.014)	(.013)		(.037)	(.037)	(.019)	(.019)	(.019)		(.049)	(.049)
Enrollment		.022	.012			.053		.041	.062			.079
		(.009)	(.026)			(.028)		(.012)	(.037)			(.036)
Enrollment squared/100			.005						−.010			
			(.011)						(.016)			
Piecewise linear trend				.136						.193		
				(.032)						(.040)		
Root MSE	6.15	6.23	6.22	7.71	6.79	7.15	8.34	8.40	8.42	9.49	8.79	9.10
N		2019		1961	471			2018		1960	471	

The unit of observation is the average score in the class. Standard errors are reported in parentheses. Standard errors were corrected for within-school correlation between classes. All estimates use f_{sc} as an instrument for class size.

Urquiola and Verhoogen (2008)

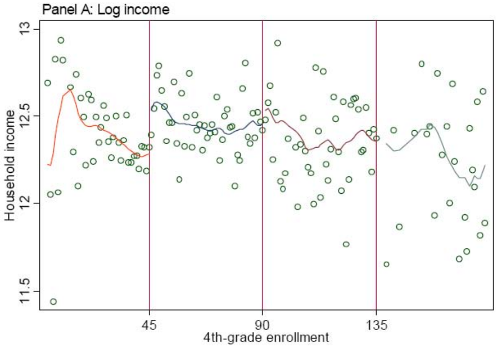
Figure 7: Histograms of 4th grade enrollment in urban private schools, 2002



Notes: Enrollment is drawn from administrative data for 2002. For visual clarity, only schools with 4th grade enrollments below 225 are displayed. This excludes less than one percent of all schools.

Urquiola and Verhoogen (2008)

Figure 8: Student characteristics and enrollment in urban private voucher schools, 2002

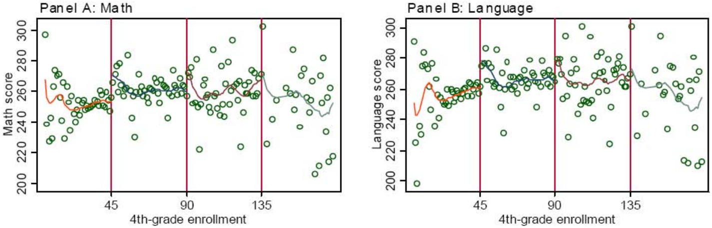


Courtesy of the American Economic Association. Used with permission.

Notes: Income and mothers' schooling come from 2002 individual-level SIMCE data aggregated to the school level. Enrollment is drawn from administrative data for the same year. The figure presents "raw" enrollment-cell means, along with the fitted values of a locally weighted regression calculated within each enrollment segment. Only data for schools with 4th grade enrollments below 180 are plotted; this excludes less than two percent of all schools.

Urquiola and Verhoogen (2008)

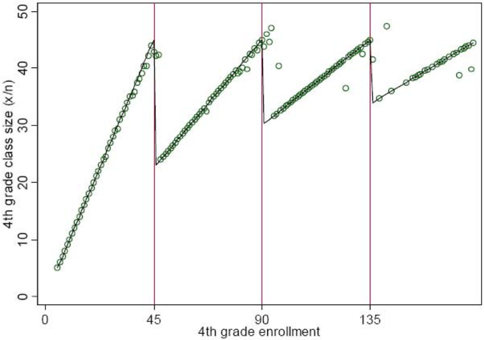
Figure 6: Test scores and enrollment in urban private voucher schools, 2002



Notes: Test scores come from 2002 individual-level SIMCE information aggregated to the school level, and enrollment is drawn from administrative data for the same year. The figures plot "raw" enrollment-cell means of test scores, along with the fitted values of a locally weighted regression calculated within each enrollment segment.

Urquiola and Verhoogen (2008)

Figure 5: 4^{th} grade enrollment and class size in urban private voucher schools, 2002



Notes: Based on administrative data for 2002. The solid line describes the relationship between enrollment and class size that would exist if the class size rule (equation 30 in the text) were applied mechanically. The circles plot actual enrollment cell means of 4^{th} grade class size. Only data for schools with 4^{th} grade enrollments below 180 are plotted; this excludes less than two percent of all schools.

Urquiola and Verhoogen (2008)

Table 5: Behavior of selected variables around enrollment cut-offs and IV specifications; urban private voucher schools, 2002

	Mothers' schooling (1)	Fathers' schooling (2)	Household income (3)	IV: Math (4)	IV: Language (5)
Class size				-0.1	0.1
				(0.1)	(0.1)
$1\{x \geq 46\}$	0.93***	0.94***	66.6***		
	(0.2)	(0.2)	(14.1)		
$1\{x \geq 91\}$	0.03	0.03	17.6		
	(0.2)	(0.2)	(17.3)		
$1\{x \geq 136\}$	0.66	0.86	143.7*		
	(0.7)	(0.8)	(79.4)		
$1\{x \geq 181\}$	0.66	0.71	53.1		
	(1.1)	(1.1)	(77.7)		
x	-0.02*	-0.02*	-2.4***	0.4***	0.4***
	(0.0)	(0.0)	(0.8)	(0.1)	(0.1)
$(x-46)*1\{x \geq 46\}$	0.02*	0.01	2.3***	-0.4***	-0.4***
	(0.0)	(0.0)	(0.8)	(0.1)	(0.1)
$(x-91)*1\{x \geq 91\}$	-0.01	0.00	-0.7	0.1	0
	(0.0)	(0.0)	(0.6)	(0.1)	(0.1)
$(x-136)*1\{x \geq 136\}$	-0.02	-0.03	-3.5	-0.2**	-0.1
	(0.0)	(0.0)	(2.3)	(0.1)	(0.1)
$(x-181)*1\{x \geq 181\}$	0.01	0.02	4	0.1	0.1
	(0.0)	(0.0)	(3.4)	(0.1)	(0.2)
Mothers' schooling				8.5***	9.5***
				(0.9)	(1.0)
Fathers' schooling				1.6*	1.1
				(0.9)	(0.9)
Household income				13.4**	16.6***
				(5.4)	(5.5)
N	1,623	1,623	1,623	1,623	1,623
R^2	0.034	0.032	0.029		

Notes: Test scores and socioeconomic status measures are from 2002 SIMCE individual-level data, aggregated to the school level. Class size and enrollment come from administrative information for the same year. *** indicates statistical significance at 1% level; ** at 5%, and * at 10%. All regressions are clustered by enrollment levels. The table focuses only on effects around the first four cut-offs, excluding the less than one percent of schools that report 4th grade enrollments in excess of 225.

Urquiola and Verhoogen (2008)

Table 3: 1st stage, reduced form, and base IV specifications; urban private voucher schools, 2002

	1st stage	Reduced form		IV	
	Class size	Math score	Language score	Math score	Language score
	(1)	(2)	(3)	(4)	(5)
Class size				-0.7***	-0.6**
				(0.3)	(0.3)
$1\{x \geq 46\}$	-16.5***	11.8***	9.9***		
	(2.7)	(3.2)	(3.3)		
$1\{x \geq 91\}$	-4.9**	0.0	1.6		
	(2.3)	(4.0)	(4.0)		
$1\{x \geq 136\}$	-4.3**	11.5	10.9		
	(2.0)	(13.6)	(12.9)		
$1\{x \geq 181\}$	-3.4	11.2	11.5		
	(3.0)	(10.6)	(13.9)		
x	0.95***	0.1	0.2*	0.8***	0.8***
	(0.01)	(0.1)	(0.1)	(0.2)	(0.3)
$(x-46)*1\{x \geq 46\}$	-0.6***	-0.1	-0.2	-0.6**	-0.6**
	(0.1)	(0.2)	(0.2)	(0.3)	(0.3)
$(x-91)*1\{x \geq 91\}$	-0.3**	0.0	-0.1	-0.2*	-0.3**
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
$(x-136)*1\{x \geq 136\}$	0.0	-0.6	-0.4	-0.2	-0.1
	(0.1)	(0.4)	(0.4)	(0.2)	(0.2)
$(x-181)*1\{x \geq 181\}$	-0.1	0.2	0.2	0.1	0.1
	(0.1)	(0.5)	(0.6)	(0.3)	(0.4)
N	1,623	1,623	1,623	1,623	1,623
R^2	0.844	0.069	0.072		

Notes: Test scores are based on 2002 SIMCE individual-level data, aggregated to the school level. Class size and enrollment come from administrative information for the same year. *** indicates statistical significance at the 1% level; ** at 5%, and * at 10%. All regressions are clustered by enrollment levels; see Lee and Card (forthcoming). The table focuses only on effects around the first four cut-offs, excluding the less than one percent of schools that report 4th grade enrollments in excess of 225 students.

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

Gender Discrimination

Esther Duflo

14.771

Outline

1. Measuring gender discrimination
2. What explains/affects gender discrimination

Measuring Gender Discrimination

Are women treated differently *because they are women*, either as a consequence of taste, statistical discrimination, or (where the word discrimination may not be pertinent) because the returns to investing in girls/women is systematically lower (independently of her abilities).

1. Differences in outcomes by gender
2. Differences in inputs by gender
3. "Experimental" Approach
4. Implicit Association Tests

Difference in outcomes by Gender: mortality

- "Missing Women" (Sen, 1986)
- Ratio of women to men in 1986: Europe: 1.05, SSA 1.022; North Africa: 0.96; China: 0.94; India 0.93
- "Missing woman": woman who is not alive but should be: number of men*1.022-number of women
- Intuition: missing women are either not born (selectively aborted) or have died earlier than men, which should reflect miss-treatment.

- Central issue with this approach (or any that look at outcomes): are there other reasons that explain why women are less likely to be born or die earlier in South Asian Countries relative to Sub-Saharian Africa.
- Two examples:
  - Are girls more likely to be born if mother has hepatitis (Oster)?(the answer turn out to be no)
  - Could the results be explained mainly by the fact that girls tend to live in larger family (due to stopping rules) (Jensen)–Jury still out on that one.
- Part of the solution to this problem will be to look at how mortality changes in responses to a situation

Other outcomes

- Education
- Labor force participation
- Wages
- Political participation

Inputs

- Expenditures on girls vs boys
- In every day life: food expenditures.
- Central problems with this approach:
  - It is difficult to measure what individual people eat in the households!
  - Girls may "need" less than boys
- Deaton: "Inferential Approach" meant to solve problem 1
  - Estimate the "cost" of a girls or a boy in term of adults goods.
  - How much money to I need to give to a household to keep their consumption of and adult good unchanged?
  - Π-Ratio : $\pi = \frac{\frac{\partial q_i}{\partial n_j}}{\frac{\partial q_i}{\partial x}}$
  - where q_i is consumption of adult good i, n_j is number of kids of gender and age group j and x is total outlay
  - ▸ Results No significant differences between girls and boys for Cote d'Ivoire (original paper)... or Pakistan (further work).
- When child is faced with a crisis: are parents less likely to hospitalize a child who is sick.

Experimental Approach

- Place people in front of two situations, otherwise similar, and ask them what they would do/how they judge the person (lab or field).
- Audit studies: how are women treated when they try to purchase a car?
- Bertrand-Mullainathan: Send resumes to firms (with black or white sounding names)

Experimental Approach (2)

- "Goldberg paradigms" experiments: place people in front of a vignette/speech and ask them to judge the person
- Beaman, Chattopadhyay, Duflo, Pande: Prejudice against female leaders in India.
- Respondent listens to a speech
- Speech is either read by a man or a woman; or vignette is presented for female pradhan name.
- Respondent must rate the speech along several dimensions
- Results People who have never been exposed to female leaders rank the speech/vignette much lower when the pradhan in the speech or vignette is a woman

Implicit Association Tests

- A method developped by psychologists (Banaji,Nosek, Greenwald, etc..) `https://implicit.harvard.edu/implicit`
- A computer based test relies on the assumption that, if a participant highly associates two concepts, she will accomplish a categorization task quicker. Detect an automatic association (implicit stereotype) by comparing response latency for different pairs of concepts
- Beaman et al. First IATs that have been done in a field setting in a low income country: used audio stimuli and a joystick
- Beaman et al. implemented three IATs: (i) association between men and women generally and concepts of good and bad; (ii) association between male and female politicians and concepts of good and bad; (iii) attitudes towards gender and domestic versus leadership activities

Implicit Association Test

- We are interested in the difference in the speed with which an individual completes a double categorization module when the classification on the screen is steroetypic versus non-steretypic
- Each test block consists of stimuli for each of the four categories represented on the screen. We randomized over whether an individual first received a stereotype or non-sterotype block
- We construct the "D-measure": difference in time taken to complete the "stereotypic block" and the "non-stereotypic" block.

Example: Male/Female Politicians and Good/Bad Stereotypic Block

Images removed due to copyright restrictions.

Left: Photograph of female politician and drawings of faces with stern expressions.

Center: Photograph of two women.

Right: Photograph of male politician and drawings of faces with pleased expressions."

IAT In Action

Photograph of people taking implicit association test (IAT) removed due to copyright restrictions.

Implicit Association Test: Stimuli

- IAT 1: Male and Female Names and Good/Bad Words
  - Good words: good, nice, fun, happiness, love, clean, sweet, heaven
  - Bad words: bad, mad, sorrow, inauspicious, sad, dirty, spoiled, hell
  - Female names: Geeta, Minoti, Anjali, Rekha, Jharna, Minu, Rupa, Basanti
  - Male names: Badal, Shyamol, Saurabh, Gopal, Anup, Ashok, Tapan, Raju
- IAT 2: Pictures of Male and Female Leaders and Good/Bad Words

Perception of Female Ability to Lead: IAT

- IAT 3: Male and Female Names and Domestic and Leadership Activities
  - Domestic activities: Eating puffed rice, Listening to the radio, Sleeping, Taking rest, Clay modelling, Harvesting, Caring for animals, Threshing rice
  - Leadership words: Gram Pradhan, Public speaking, Govern, Chairperson, Lobbying, Leader, Campaigning, Taking bribes
  - Pictures of Men and women are shown doing activities which should invoke the concept of politicians and leaders, such as speaking on a microphone or in front of a group

Example: Gender and Leadership/Domestic Activities

Stereotypic Block

Images removed due to copyright restrictions.
Left: photographs of woman and outdoor scene.
Center: speech bubble with the name Ashok.
Right: photographs of man and building."

Results

What explains (and affects) gender discrimination?

- Extreme situation for the household
  - Back to Das Gupta and Ray model: when things are very tight, households may need to focus resources on one child, may be preferably a boy.
  - Rose: Relative increase in girls' mortality in period of drought, especially for landless households ▸ Table

Differential Returns to investment

- Foster and Rosenzweig (2001): when there is greater technological progress outside the village, we observe relative improvement in girls survival rate, the opposite is true if there is technological project inside the village ▸ Table
- Qian: tea in China
  - Women have comparative advantage in tea
  - Men have comparative advantage in orchard
  - When households were given the right to cultivate cash crops, in counties that produced tea, relative girls mortality improved; in countries that produced orchard, it worsened ▸ Graph ▸ Table
  - Explanation: it could be either differential returns (investment) or better bargaining power for the mother. Difficult to distinguish

Exposure

- Persistence of discrimination against women (especially in education; jobs; leadership) may be the result of lack of exposure: people may statistically discriminate against women because they have never seen women in action.
- Beaman et al see how measure of discrimination (IAT, responses to vignettes) change after voters have been exposed to a women due to a mandated represenation program
- Village council randomly assigned to be “reserved” for female leaders
- In “reserved” villages, only women can be elected as village head.
- Results: ▸ Taste IAT does not change with exposure. But statistical discrimination is completely erased for men (and not for women) ▸ Stereotypical IAT ▸ Speech and Vignettes

Deaton (1989)

Outlay Equivalence Ratios for Adult Goods, Côte d'Ivoire, 1985

Gender and age[a]	*Adult clothing*	*Adult fabric*	*Adult shoes*	*Alcohol*	*Tobacco*	*Meals out*	*Entertainment*	*All adult goods*
				π-ratios				
Children								
M 0-4	-0.01	0.37	-0.23	-0.58	-0.89	0.27	-0.42	-0.12
M 5-14	-0.67	-0.35	-0.21	-0.69	-0.71	-0.37	-0.41	-0.49
F 0-4	-0.20	0.41	-0.24	-0.33	-0.45	-0.21	-0.30	-0.22
F 5-14	-0.39	-0.12	-0.45	-0.39	-0.45	-0.62	-0.37	-0.48
Adults								
M 15-55	1.30	0.79	1.63	-0.19	1.88	0.91	0.74	0.81
M > 55	-0.74	0.33	-0.28	1.45	1.06	-0.47	-0.98	0.16
F 15-55	0.32	-0.14	0.17	-0.39	-0.41	-1.33	-1.07	-0.71
F > 55	-1.14	-0.97	-0.69	-1.24	-1.29	-1.33	-0.99	-1.21
				Standard errors				
M 0-4	0.46	0.35	0.33	0.38	0.74	0.32	0.40	0.20
M 5-14	0.32	0.25	0.23	0.27	0.52	0.22	0.28	0.14
F 0-4	0.43	0.33	0.30	0.36	0.69	0.30	0.37	0.19
F 5-14	0.33	0.25	0.23	0.27	0.53	0.23	0.29	0.14
M 15-55	0.31	0.22	0.23	0.22	0.54	0.20	0.25	0.13
M > 55	0.45	0.35	0.32	0.39	0.75	0.31	0.39	0.20
F 15-55	0.27	0.20	0.19	0.22	0.43	0.19	0.23	0.11
F > 55	0.53	0.40	0.37	0.44	0.85	0.37	0.46	0.23

a. M = Male; F = Female.
Source: Deaton (1987).

Figure by MIT OpenCourseWare.

	Average Coefficients							
			Pradhan					
	Average Effect		Perform Duties Well		Is Effective		Cares about Villagers' Welfare	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
anel A								
Female Pradhan	-0.055	-0.035	-0.076	-0.043	-0.047	-0.029	-0.055	-0.003
	(0.027)	(0.031)	(0.036)	(0.042)	(0.032)	(0.034)	(0.033)	(0.032)
Female Pradhan * Ever Reserved	0.096	0.020	0.121	0.026	0.084	0.022	0.102	0.008
	(0.037)	(0.039)	(0.049)	(0.051)	(0.042)	(0.043)	(0.043)	(0.041)
Test: Female Pradhan + Female Pradhan * Ever Reserved	0.041	-0.014	0.045	-0.018	0.037	-0.006	0.047	0.005
	(0.024)	(0.023)	(0.032)	(0.029)	(0.027)	(0.026)	(0.027)	(0.025)

Implicit Associations

Rose (1999)

Logits Dependent Variable: Girl

Sample	*Pooled*		*Landed*		*Landless*	
	All (1)	*No floods (2)*	*All (3)*	*No floods (4)*	*All (5)*	*No floods (6)*
RF shock (year of birth)	0.26 (2.6)	0.39 (2.4)	-0.02 (0.1)	0.17 (0.9)	0.89 (3.5)	0.92 (2.0)
RF shock (age 1)	0.39 (2.6)	0.57 (3.5)	0.40 (2.2)	0.54 (2.5)	0.11 (0.23)	0.63 (1.5)
RF shock (age 2)	0.27 (1.5)	0.09 (0.4)	0.09 (0.4)	-0.22 (0.9)	0.73 (1.8)	0.75 (1.4)
RF shock (age 3)	0.11 (0.7)	0.03 (0.2)	0.10 (0.5)	-0.06 (0.3)	0.70 (1.8)	0.60 (1.4)
RF shock (age 4)	0.20 (1.0)	0.29 (1.4)	0.28 (1.3)	0.13 (0.5)	0.33 (1.0)	0.59 (1.4)
RF shock (age 5)	0.11 (0.8)	0.17 (1.1)	0.11 (0.6)	0.06 (0.3)	0.13 (0.4)	0.40 (1.2)
Land owned	0.001 (1.5)	0.002 (1.8)	0.002 (1.5)	0.002 (1.9)		
Mother educated	0.25 (2.1)	0.31 (2.1)	0.23 (1.5)	0.27 (1.6)	0.39 (1.5)	0.44 (1.3)
Head educated	-0.28 (2.5)	-0.38 (3.3)	-0.32 (2.3)	-0.50 (3.2)	-0.16 (0.7)	-0.16 (0.7)
Educational institution	-0.28 (2.5)	-0.30 (2.5)	-0.13 (1.1)	-0.17 (1.3)	-1.2 (3.4)	-0.99 (3.0)
Health institution	0.19 (1.6)	0.21 (1.6)	0.14 (0.9)	0.22 (1.3)	0.38 (2.1)	0.25 (1.2)
Factor1	0.25	0.25	0.25	0.25	0.25	0.25
Factor2	0.83	0.82	0.82	0.80	0.85	0.86
N	2,297	1,926	1,722	1,418	575	508

Figure by MIT OpenCourseWare.

Foster and Rosenzweig (2001)

Determinants of the Difference in Mortality Rates of Boys and Girls: Children Aged 0-4

Variable/Estimation procedure	*OLS: 1971*	*FE-IV: 1971-82*	*FE-IV: 1971-82*	*FE-IV: 1971-82*
Log of land price - village	.0944 (1.82)	.00238 (0.08)	-.373 (2.37)	-.415 (2.41)
Mean log of land price - marriage market (Radius=67Km)	–	–	.343 (2.09)	.434 (2.20)
Mean log of land price - marriage market (Radius>67, <314Km)	–	–	–	–
Mean log of land price - marriage market (Radius>314, <1000Km)	–	–	–	–
Mean log of yield - village	-.0292 (1.60)	-.0105 (0.36)	.00794 (0.08)	.0364 (0.32)
Mean log of yield - marriage market	–	–	-.0115 (0.10)	-.0466 (0.36)
Mean household wealth ($x10^{-5}$) - village	-.0838 (1.48)	.00740 (0.10)	.104 (0.98)	.197 (1.21)
Mean household wealth ($x10^{-6}$) - marriage market	–	–	–	-.247 (1.04)
Proportion mother literate - village	.0918 (1.94)	.181 (2.79)	.0418 (0.38)	.0312 (0.22)
Proportion mother literate - marriage market	–	–	–	-.0169 (0.11)

Figure by MIT OpenCourseWare.

Qian (2008)

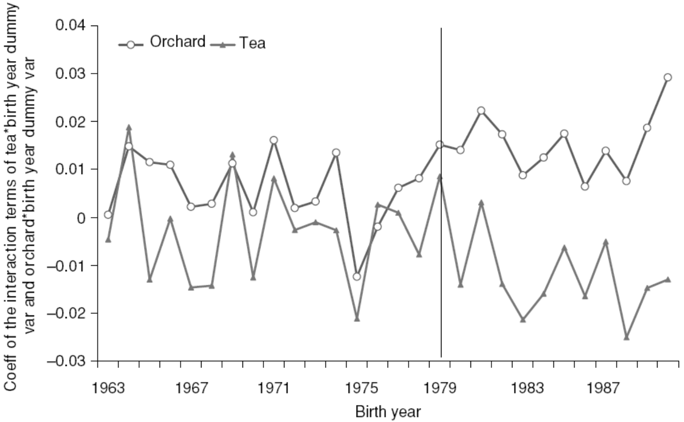


FIGURE V

The Effect of Planting Tea and Orchards on Sex Ratios

Coefficients of the interactions of birth year × amount of tea planted and birth year × amount of orchards planted controlling for year and county of birth FEs.

Qian (2008)

TABLE III
OLS AND 2SLS ESTIMATES OF THE EFFECT OF PLANTING TEA AND ORCHARDS ON SEX RATIOS CONTROLLING FOR COUNTY LEVEL LINEAR COHORT TRENDS

	Dependent variables					
	Fraction of males			Tea × post	Fraction of males	
	(1) OLS	(2) OLS	(3) OLS	(4) 1st	(5) IV	(6) IV
Tea × post	−0.012	−0.013	−0.012		−0.072	−0.011
	(0.007)	(0.006)	(0.005)		(0.031)	(0.007)
Orchard × post	0.005					
	(0.002)					
Slope × post	−0.002			0.26		
	(0.002)			(0.057)		
Linear trend	No	No	Yes	Yes	No	Yes
Observations	28,349	37,756	37,756	37,756	37,756	37,756

Notes. Coefficients of the interactions between dummies indicating whether a cohort was born post-reform and the amount of tea planted in the county of birth. All regressions include county and birth year fixed effects and controls for Han, and cashcrop × post. All standard errors are clustered at the county level. In column (1), the sample includes all individuals born during 1970–1986. In columns (2)–(6), the sample includes all individuals born during 1962–1990. Post = 1 if birthyear > 1979. Data for land area sown are from the 1997 China Agricultural Census.

Table 3: Implicit Association tests

	Male/Female Names and Good/Bad IAT		Male/Female Politician and Good/Bad IAT	
	Male	Female	Male	Female
	(3)	(4)	(5)	(6)
Panel A				
Ever Reserved	-0.001	0.006	-0.008	-0.015
	(0.032)	(0.042)	(0.034)	(0.037)
Panel B				
First Reserved 2003	-0.039	0.020	-0.005	0.010
	(0.042)	(0.051)	(0.049)	(0.049)
Reserved 1998 and 2003	0.039	0.044	0.004	-0.008
	(0.041)	(0.068)	(0.052)	(0.052)
Only Reserved 1998	0.011	-0.048	-0.020	-0.043
	(0.047)	(0.051)	(0.044)	(0.051)
Test: 2003 = both 1998 and 2003 = 1998	0.301	0.299	0.908	0.636
Mean of Never Reserved Sample	0.134	-0.157	0.093	-0.079
	(0.025)	(0.026)	(0.027)	(0.025)
N	510	408	554	510

Notes:

Table 4: Implicit Association Test Measure of Implicit Bias and Ladder

	Leadership/Domestic and Male/Female IAT	
	Male	Female
	(7)	(8)
Panel A		
Ever Reserved	-0.070	0.022
	(0.030)	(0.041)
Panel B		
First Reserved 2003	-0.089	0.104
	(0.040)	(0.053)
Reserved 1998 and 2003	-0.024	-0.079
	(0.045)	(0.067)
Only Reserved 1998	-0.080	-0.021
	(0.039)	(0.050)
Test: 2003 = both 1998 and 2003 = 1998	0.390	0.032
Mean of Never Reserved Sample	0.110	0.150
	(0.021)	(0.027)
N	477	357

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

14.771: Public Finance Lecture 1

Ben Olken

October 2008

Outline

- Topics I will talk briefly about:
  - Tax (today)
  - Market failures (today)
    - Correcting externalities
    - Providing public goods
  - Redistribution (next lecture)
- Topics I won't talk about at all:
  - Social insurance
  - Regulation
  - Federalism, decentralization, and local public goods
  - (among many other things)
- Note: Very little work in this area. Good area for research!

Tax

- There is a vast literature in PF on taxation. E.g., incidence, optimal income tax theory, capital taxation, consumption taxes, dynamic considerations, etc, etc.
- By comparison we know very little about tax – either theory or empirics – in developing countries.
- What we do know suggests that there is a fundamental difference between developing and developed countries:
  - *Information.* There is much less information available. How do you levy an income tax on people who are subsistence farmers? Or laborers in an all-cash economy?
  - *Enforcement.* Given the information problems there is substantial opportunity for corruption.

Tax

- As a result of information and enforcement problems, tax the tax structure in developing countries looks very different than in developed countries, because you need to tax things with high information and low elasticites of evasion (Gordon and Li 2005)

Developed and developing countries tax structure

GDP per capita	*Tax revenue (% of GDP)*	*Income taxes (% of revenue)*	*Corporate income tax (% of income taxes)*	*Consumption and production taxes (% of revenue)*	*Border taxes (% of revenue)*	*Inflation rate*	*Seignorage income (% of revenue)*	*Informal economy (% of GDP)*
< $745	14.1	35.9	53.7	43.5	16.4	10.6	21.8	26.4
$746-2,975	16.7	31.5	49.1	51.8	9.3	15.7	24.9	29.5
$2,976-9,205	20.2	29.4	30.3	53.1	5.4	7.4	6.0	32.5
All developing	17.6	31.2	42.3	51.2	8.6	11.8	16.3	30.1
> $9,206	25.0	54.3	17.8	32.9	0.7	2.2	1.7	14.0

Figure by MIT OpenCourseWare.

from Gordon and Li (2005)

Tax

- As a result of information and enforcement problems, tax the tax structure in developing countries looks very different than in developed countries, because you need to tax things with high information and low elasticites of evasion (Gordon and Li 2005)
  - Smaller: 2/3 the size of tax revenue in rich countries as percentage of GDP
  - Income taxes focus on corporate, not individual.
  - Tariffs and seigniorage play non-trivial role much more important

Explanation

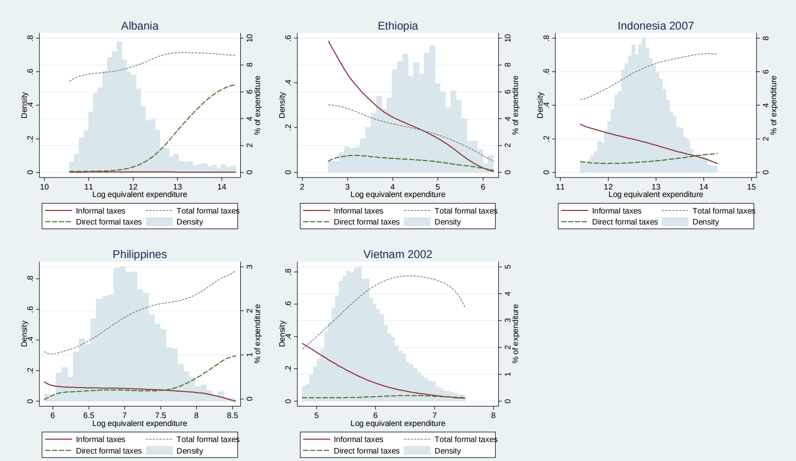
- One explanation: information (Gordon and Li 2005).
  - Using the financial sector generates information for the government.
  - Taxes focus on corporate because the large corporations are inelastic in their use of the formal banking system, so this is where taxes are focused.
  - Tariffs protect the taxed sector.
  - Inflation taxes the cash economy.
- Seems intuitive, but far from the last word on the subject.

Informal taxation

- Olken and Singhal (2008) study phenomenon of 'voluntary' contributions to local public goods
  - *Harambee* in Kenya
  - *Gotong Royong* in Indonesia
- Idea: taxation analogue of informal insurance
  - Specifically, local communities have good information about incomes, but face enforcement constraints
  - They can therefore enforce 'voluntary' contributions to public goods – what we call informal taxation – through social sanctions
  - Within communities, rich pay more, but less as a share of expenditure, so it is regressive
  - Social sanctions less potent in richer, urban areas so this is primarily a rural phenomenon
  - On net: makes tax system more regressive

Informal taxation

- To examine who pays different types of taxes, we run Fan locally-weighted regressions of taxes as share of expenditure against expenditure per equivalent adult



How big is evasion?

- Fisman and Wei (2004): what is the 'elasticity' of tax evasion with respect to tax rates?
- Empirical challenge: very hard to measure what the true tax assessment should be.
- Fisman and Wei's idea:
  - Look at both sides of the China - Hong Kong border, where China is the 'high evasion' side and Hong Kong is the 'low evasion side'
  - Denote the difference between what Hong Kong (low corruption) and China (high corruption) reports as evasion, i.e,

$$gap\_value = \log(\text{export\_value}) - \log(\text{import\_value})$$

- Key regressions:

$$gap\_value_k = \alpha + \beta_1 tax_k + \varepsilon_k$$

$$gap\_value_k = \alpha + \beta_1 tax_k + \beta_2 tax\_o_k + \varepsilon_k$$

Results

Effect of Tax Rates on Evasion (Measured in Value)

	Regression						
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Tax rate	2.93 (.74)	2.46 (.67)	3.21 (.87)	3.57 (.89)	2.98 (.81)	2.61 (.79)	3.4 (.96)
Constant	-1.31 (.29)	-1.04 (.23)	-1.31 (.30)	-1.48 (.31)	-1.29 (.29)	-1.12 (.27)	-1.46 (.34)
Excluding outliers?	no	yes	no	no	yes	yes	yes
Excluding products lacking tax on similar products?	no	no	yes	no	no	yes	yes
Excluding products lacking observations on quantities?	no	no	no	yes	yes	no	yes
Observations	1,663	1,639	1,470	1,102	1,087	1,450	968
R^2	.020	.017	.022	.031	.025	.017	.029

Note: The dependent variable is log (value of exports from Hong Kong to China) - log (value of imports to China from Hong Kong). Robust standard errors are in parentheses, accounting for clustering of standard errors by four-digit HS.

Figure by MIT OpenCourseWare.

- $\beta_1 = 3$: One percentage point increase in taxes on your product increase evasion gap by 3%

Results

Incorporating the Average Tax on Similar Products
Dependent Variable: Log (Value of Exports from Hong Kong to China) - Log (Value of Imports to China from Hong Kong)

	Regression				
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
Tax rate		6.07 (1.37)	5.31 (1.25)	8.32 (1.56)	7.46 (1.42)
Tax on similar products	2.62 (.90)	-3.16 (1.39)	-2.98 (1.33)	-4.65 (1.58)	-4.45 (1.53)
Constant	-1.09 (.034)	-1.20 (.31)	-1.02 (.28)	-1.56 (.38)	-1.33 (.35)
Excluding outliers?	no	no	yes	no	yes
Excluding products lacking observations on quantities?	no	no	no	yes	yes
Observations	1,470	1,470	1,450	981	968
R^2	.014	.025	.020	.041	.035

Note: Robust standard errors are in parentheses, accounting for clustering of standard errors by four-digit HS.

Figure by MIT OpenCourseWare.

- $\beta_1 = 6, \beta_2 = -3$: Less evasion when nearby products also have higher tax rates implies reclassification is an important mechanism

Empirical issues to think about

- Are tax rates endogenous?
  - They should be: governments should understand that there is differential elasticities and set lower tax rates on more elastic (easier to evade) items.
  - This paper assumes government is naiive.
  - How would sophisticated government bias results?
- Is spillover model mis-specificied?
  - Recall they include all products in category, but products vary in size.
  - Suppose in a category you have one very large product and very small product.
  - Small shift out of large product is a large proportional shift into small product.
  - Might have been better to include only one product per category? Or model this explicitly?

Tax

- That's it!
- (Not quite, but you get the point)

Correcting externalities

- Basic public finance theory: use taxes and subsidies to correct externalities.
  - These taxes and subsidies are permanent: as long as the externality remains, you need the tax/subsidy in place to correct it.
  - Optimal tax/subsidy depends on degree of externality, elasticity of supplying externality, and marginal cost of public funds
- Many in development design policies for 'sustainability.'
  - Idea is that there are multiple equilibria, so a one time intervention can lead to a 'sustainable' outcome.
- Is this plausible?

The Illusion of Sustainability

- Kremer and Miguel (2007)
- Setting:
  - Deworming in Kenya.
  - Disease transmission implies very large positive externalities from deworming
- Research design: randomized experiment to
  - Test elasticity of demand for deworming drugs by introducing cost-sharing (at approx. 20% of actual average cost)
  - Test whether social adoption spillovers are positive (this could help generate multiple equilibria)

Experimental Design

- Kremer and Miguel (2007) experimental design:

Kremer and Miguel (2007) basic experimental design

	1998	1999	2000	2001
Group 1 (25 schools)	Free Treatment	Free Treatment	Free Treatment	Free Treatment Cost-Sharing
Group 2 (25 schools)	Control	Free Treatment	Free Treatment	Free Treatment Cost-Sharing
Group 3 (25 schools)	Control	Control	Control	Free Treatment

- Examine Groups 2 and 3 in 2001. Collect data on usage and average number of links to households in different experimental treatments
- Estimate Probit model with errors clustered by school:

$$P\left(T_{ij}=1\right)=\Phi\left(N_{ij}^{E}a+N_{ij}^{\prime}b_{1}+b_{2}COST_{j}+Z_{ij}^{\prime}b_{3}\right)$$

Results on cost sharing

TABLE VII
THE IMPACT OF COST-SHARING

	Dependent variable: Child took deworming drugs in 2001		
	(1)	(2)	(3)
Explanatory variables:			
Cost-sharing school indicator	−0.580***	−0.459***	−0.572***
	(0.054)	(0.122)	(0.080)
Cost-sharing *Respondent years of education		0.002	
		(0.007)	
Cost-sharing *Community group member		0.021	
		(0.072)	
Cost-sharing *Total number of children		−0.021	
		(0.016)	
Cost-sharing *Iron roof at home		−0.047	
		(0.064)	
Effective price of deworming per child(= cost/# household children in that school)			−0.001
			(0.002)
1/(# household children in that school)			−0.348***
			(0.066)
Social links, other controls	Yes	Yes	Yes
Number of observations (parents)	1,678	1,678	1,678
Mean of dependent variable	0.61	0.61	0.61

Cost calculation

- Cost per-pupil under full subsidy: US \$1.478
- Cost per-pupil under cost sharing US \$1.374
    - Assumes \$15 per school fixed cost, US\$0.03 marginal cost to collect funds, and US\$0.30 cost sharing
- Assume fixed budget B
- Extra students treated with cost-sharing:

$$\frac{B}{1.374} - \frac{B}{1.478} = B * 0.0512$$

- Extra revenue collected from cost-sharing:

$$\frac{B * US\$0.30}{1.374} = B * 0.2183$$

- Marginal social cost of additional student treated:

$$\frac{B * 0.2183}{B * 0.0512} = US\$4.26$$

- Marginal cost of public funds would need to be greater than 4.26!

Results on social spillovers

Explanatory variables:		
# parent links with children in early treatment schools (Groups 1 and 2, not own school)	−0.031** (0.014)	−0.040** (0.017)
# parent links with children in early treatment schools * Group 2 school indicator		0.017 (0.029)
Proportion direct (first-order) parent links with children in early treatment schools		
# parent links with children in early treatment schools, with whom respondent speaks at least twice/week		
# parent links with children in early treatment schools, with whom respondent speaks less than twice/week		
# parent links with children in Group 1, 2, or 3 schools, not own school, with whom respondent speaks at least twice/week		
# parent links with children in Group 1, 2, or 3 schools, not own school, with whom respondent speaks less than twice/week		
# parent links with children in early treatment schools * Respondent years of education		
# parent links with children in Group 1, 2, or 3 schools, not own school	0.013 (0.011)	0.012 (0.017)

Implications

- Very price elastic: introducing cost sharing dramatically reduces take-up
  - Suggests probably more effectively financed out of general public funds than out of user fees
  - Social spillovers (in this case) are negative, which leads in direction of single equilibrium
- Consistent with many other findings of very high price-elasticity for health (including those discussed by Esther):
  - Cohen and Dupas (2007)
  - Banerjee et al. (2008)

Public goods

- Public good definition: non-rival, non-excludable goods
- In practice, government also provides other types of goods with large fixed costs:
  - Dams
  - Electric power
  - Airports
- Issues to think about in developing countries:
  - Return and willingness to pay – does social return exceed social cost?
  - Distributional impacts?
  - What level of government should provide? Does Tiebout sorting model apply in developing countries?
  - Corruption?

Dams

- Duflo and Pande (2007)
- Setting:
  - Irrigation dams provide irrigation downstream, impose costs upstream.
- Key questions:
  - What is overall effect of the dams?
  - Do transfers happen?
- Empirical idea:
  - Gradient of river determines whether a site is a good candidate for a dam. Intermediate slope best for irrigation dams,
  - Use national trends in dam construction, times initial share of dams in state, to estimate number of new dams built in state in a year.
  - Instrument is the interaction of predicted dams built in state with river gradient in district to get 'predicted new dams in district'
  - Is exclusion restriction valid?

Results

	Per-capita expenditure
	(1)
Dams	
Own district	−0.289
	(0.115)
Upstream	0.093
	(0.057)
Dams	
Own district	−0.457
	(0.467)
Upstream	0.142
	(0.084)
N	1,799
First stage F-statistic (own district)	7.71

Results

- Positive effects on agriculture in downstream districts, (potentially) negative effects in district with dam
- Reductions in poverty in downstream districts, increases in poverty in district with dam
    - So transfers are not happening
- In fact, increases in poverty in district with dam exceed reductions in poverty in downstream districts
    - Reductions in log per-capita expenditure are 0.289 OLS, 0.457 IV
    - Increases only 0.093 OLS, 0.142 IV. 1.75 times as many downstream as upstream, so average log per-capita expenditure goes down.
    - Also note log dependent variable makes 'adding up' hard to do, so transfers might still be possible.
- Note: are these estimates too large?

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

14.771: Public Finance Lecture 2

Ben Olken

October 2008

Outline

- Basic problem: lack of information about who is really poor.
  - This is a problem everywhere. See US PF literature.
  - But the problem is particularly severe in developing countries: we don't even observe income!
- Two approaches:
  - Broad subsidies (e.g., food subsidies)
  - Try to do targeted transfers anyway

Poverty metrics

- Standard decomposable metric developed by Foster, Greer, and Thorbecke (1984):
  - Define z as the poverty line.
  - Then for $\alpha \geq 0$ define

$$P_\alpha = \int_0^z \left(\frac{z-y}{z}\right)^\alpha f(y)\, dy$$

- Special cases:
  - $P_0 = \int_0^z f(y)\, dy$ is the "headcount" ratio, i.e., number of poor people
  - $P_1 = \int_0^z \left(\frac{z-y}{z}\right) f(y)\, dy$ is the "poverty gap",i.e., the amount of money required to bring all poor people up to the poverty line.
  - $\alpha > 1$ puts more weight on the poverty of very poor.
- Key property is decomposability. Assume i subgroups with population shares λ_i. Then

$$P_\alpha = \sum_i \lambda_i P_{i,\alpha}$$

Thinking about transfers

- Assume for the moment we cannot directly identify poor households (i.e., no targeting)
- Besley and Kanbur (1988): How do we evaluate subsidies in terms of poverty reductions?
  - Infra-marginal subsidies
    - To everyone
    - With geographical targeting
  - Marginal subsidies (i.e., price changes)
    - To everyone
    - When there are both producers and consumers

Notation

- Since we're talking about subsidies we sometimes need two price vectors:
  - p is the undistorted world price vector
  - q is the price vector faced by households
- Indirect utility function: $V(q, y)$
- Define equivalent income as income at world price vector, i.e.

$$y_E(p, q, y),$$

defined by

$$V(p, y_E) = V(q, y)$$

Infra-marginal subsidies

- Typically happen in the form of ration shops, where each household entitled to buy x kg of subsidized food
- Can be thought of as lump-sum transfer of size m, where m is monetary equivalent of subsidy at p prices
- Impact on poverty:

$$P_\alpha = \int_0^z \left[\frac{z_E - y_E(p, p, y+m)}{z_E} \right]^\alpha f(y)\, dy$$

- Taking derivatives with respect to m:

$$\begin{aligned} \frac{\partial P_\alpha}{\partial m} &= \frac{\alpha}{z_E} \int_0^z \left[\frac{z_E - y_E(p, p, y+m)}{z_E} \right]^{\alpha-1} \left(-\frac{\partial y_E}{\partial m} \right) f(y)\, dy \\ &= -\frac{\alpha}{z_E} P_{\alpha-1} \end{aligned}$$

- So if we care about poverty gap ($\alpha = 1$), then impact of inframarginal subsidy is proportional to the headcount ratio.

Geographical targeting

- Geographical targeting is much easier than individual targeting, since we can use representative household surveys to figure out the geographical distribution of poverty
- This allows us to improve substantially on lump-sum transfers.
- Suppose i regions, population shares represented by λ_i.
- Increasing budget to region i by b_i gives each person in region i a transfer of $\frac{b_i}{\lambda i}$
- Using the logic from before,

$$\frac{\partial P_\alpha}{\partial b_i} = -\frac{\alpha}{z_E} P_{i,\alpha-1}$$

- So, if objective is to minimize national P_α, give infra-marginal subsidies at the margin to regions with highest $P_{\alpha-1}$. I.e., to reduce poverty gap, put ration shops in areas with high poverty rates, since that is where money most efficiently reaches the poor.

Geographical targeting

- How is geographical targeting done in practice?
- One approach: Poverty maps. Elbers, Lanjouw and Lanjouw (2003)
- Idea:
  - Representative household survey has data on consumption, for small number of people
  - Census has data on every individual (age, education, etc), but doesn't measure consumption
  - So project consumption on census characteristics in household survey, and use census to extrapolate out of sample
  - Standard errors need to be corrected for spatial autocorrelation
- Big savings in cost:
  - In Cambodia, geographic targeting at province level reduces cost of given poverty reduction by 45%; targeting at commune level reduces cost of given poverty reduction by 69%! (Elbers et al 2007)

Price subsidies at the margin

- Price subsidies also affect consumer choices.
- Notation:
  - Post-tax prices: $q_i = p_i + t_i$
  - Effect on poverty of change in subsidy t_i:

$$\frac{\partial P_\alpha}{\partial t_i} = \frac{\alpha}{z_E} \int_0^z \left[\frac{z_E - y_E(p, q, y)}{z_E} \right]^{\alpha - 1} \left(-\frac{\partial y_E}{\partial q_i} \right) f(y)\, dy$$

  - Consumer demand $x_i(q, y)$. Define

$$\begin{aligned} \overline{x}_i &= \int_0^\infty x_i f(y)\, dy \text{ (mean consumption of } i\text{)} \\ \overline{x}_i^P &= \frac{\int_0^z x_i f(y)\, dy}{\int_0^z f(y)\, dy} \text{ (mean consumption of } i \text{ by poor)} \end{aligned}$$

  - Government budget constraint:

$$\int_0^\infty \left[\sum_k t_k x_k(q, y) \right] f(y)\, dy = B$$

Effect of a revenue-neutral change in taxes

- Consider taxes on two commodities, t_1 and t_2.
- Budget balance implies

$$\frac{dt_1}{dt_2} = \frac{\int_0^\infty \left(\sum_k t_k \frac{\partial x_k}{\partial t_2} + x_2\right) f(y)\, dy}{\int_0^\infty \left(\sum_k t_k \frac{\partial x_k}{\partial t_1} + x_1\right) f(y)\, dy}$$

- Effect of budget-neutral increase in t_1 is:

$$\frac{\partial P_\alpha}{\partial t_1} = \frac{\alpha}{z_E} \int_0^z \left[\frac{z_E - y_E}{z_E}\right]^{\alpha-1} \left(-\frac{\partial y_E}{\partial q_1} - \frac{\partial y_E}{\partial q_2}\frac{dt_2}{dt_1}\right) f(y)\, dy$$

- To gain intuition, need to understand how equivalent income affected by subsidies, i.e., $\frac{\partial y_E}{\partial q_i}$.

Effect of a revenue-neutral change in taxes

- Simple case: suppose we start from case of no subsidies, so $t_k = 0$ $\forall k$. Then (recalling 14.121)

$$\left.\frac{\partial y_E}{\partial q_i}\right|_{p=q} = -x_i\left(q, y\right)$$

$$\begin{aligned}
\frac{\partial P_\alpha}{\partial t_1} &= \frac{\alpha}{z_E}\int_0^z \left[\frac{z_E - y_E}{z_E}\right]^{\alpha-1}\left(-\frac{\partial y_E}{\partial q_1} - \frac{\partial y_E}{\partial q_2}\frac{dt_2}{dt_1}\right) f\left(y\right) dy \\
&= \frac{\alpha}{z_E}\int_0^z \left[\frac{z_E - y_E}{z_E}\right]^{\alpha-1}\left(x_1 + x_2\frac{dt_2}{dt_1}\right) f\left(y\right) dy \\
&= \frac{\alpha}{z_E}\int_0^z \left[\frac{z_E - y_E}{z_E}\right]^{\alpha-1}\left(x_1 - x_2\frac{\overline{x}_1}{\overline{x}_2}\right) f\left(y\right) dy \\
&= \frac{\alpha}{z_E\overline{x}_1}\int_0^z \left[\frac{z_E - y_E}{z_E}\right]^{\alpha-1}\left(\frac{x_1}{\overline{x}_1} - \frac{x_2}{\overline{x}_2}\right) f\left(y\right) dy
\end{aligned}$$

- Reduction in P depends on relative consumption of x_1 and x_2 by poor

Effect of a revenue-neutral change in taxes

- Special case of $\alpha = 1$ (poverty gap).
- Define H as headcount ratio (fraction poor). Then:

$$\begin{aligned}\frac{\partial P_\alpha}{\partial t_1} &= \frac{\alpha}{z_E \overline{x}_1} \int_0^z \left(\frac{x_1}{\overline{x}_1} - \frac{x_2}{\overline{x}_2} \right) f(y)\, dy \\ &= \frac{\alpha}{z_E \overline{x}_1} H \left(\frac{\overline{x}_1^P}{\overline{x}_1} - \frac{\overline{x}_2^P}{\overline{x}_2} \right)\end{aligned}$$

- Very intuitive: subsidize the commodity where share of commodity consumed by the poor is highest, if goal is to reduce P_1.
- More generalized versions have similar intuitions with appropriate weights.
- If initial taxes not equal to 0, also need to incorporate effect of tax change on other revenues

Infra-marginal vs. marginal subsidies

- Assume positive Engel curves on all goods, so expenditure on all goods increases with income.
  - Then infra-marginal subsidies are always better than marginal subsidies.
  - Intuition: for marginal subsidies, effect on poverty only from share of expenditure from the poor, $\overline{x}_1^P$

Producers and consumers

- Assume income generated by profit function

$$y = \Pi\left[q, k\right]$$

where k are endowments like land.

- For producers,

$$\left.\frac{\partial y_E}{\partial q_i}\right|_{p=q} = -\left[x_i\left(q, y\right) - r_i\left(q, k\right)\right]$$

where r is production of commodity. (envelope theorem).

- Define $n = r - x$.
- Then effect of price change is

$$\begin{aligned}\left.\frac{\partial P_\alpha}{\partial t}\right|_{p=q} &= \lambda_1 \frac{\alpha}{z_E}\int_0^z \left[\frac{z_E - y_E}{z_E}\right]^{\alpha-1} x f_1\left(y\right) dy + \\ &\quad \lambda_2 \frac{\alpha}{z_E}\int_0^z \left[\frac{z_E - y_E}{z_E}\right]^{\alpha-1} n f_2\left(y\right) dy\end{aligned}$$

Producers and consumers

- If $\alpha = 1$, this simplifies to

$$\left.\frac{\partial P_{\alpha}}{\partial t}\right|_{p=q} = \frac{\alpha}{z_E}\left(\lambda_1 H_1 \overline{x}_1^P + \lambda_2 H_2 \overline{n}_2^P\right)$$

- This is intuitive: effect on poverty depends on mean net consumption among consumers and mean net consumption among producers.

Summary so far

- Inframarginal subsidies tend to be better than price subsidies, unless there are inferior goods that you can subsidize.
- Why?
  - Higher share goes to the poor
  - Don't hurt producers
  - Can do even better with geographic targeting
  - Also: dead-weight loss from distorted prices
- But inframarginal subsidies are much harder to implement (e.g., corruption, operating shops, etc)
- And, even they are not perfect, because large amounts of transfers still go to non-poor.
- Can we do better with more directly targeted transfers?

Targeting

- Targeting options if income is not observable:
  - Proxy-means tests
  - Self-targeting
  - Community-based targeting

Proxy-Means Tests

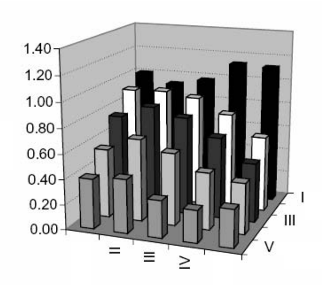
- Similar idea to poverty mapping, but at individual level. This is the main way individual targeting is done in most developing countries. (E.g, Progresa).
- Concept: consumption surveys are expensive, and non-verifiable, so you can't use them to target directly
- Instead: do a survey where you collect data on assets (land, house, motorcycle, etc)
  - Assets capture permanent component of income
  - And they are hard to falsify on a survey
- Use survey data to estimate relationship between consumption and assets, and used predicted consumption for targeting
- Problems
  - R^2 much less than 1, so you don't get poverty exactly right
  - Corruption among surveyors
  - Costly: need to do a census

Self-Targeting

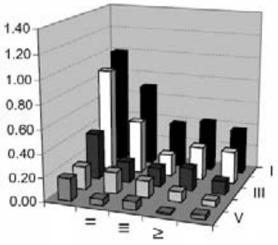
- Nichols and Zeckhauser (1982): "Ordeals" can be used to target the poor
  - Suppose you need to wait in long line to get unemployment benefits
  - Unemployed have low opportunity cost of time, so they are more likely to wait in line
  - Waiting in line therefore serves as a screening device

Self-Targeting In Practice

- Sumarto et al (2003) compares targeting of two programs in Indonesia in 1998
  - Subsidized rice (no self targeting)
  - Public employment scheme (self targeting)



Rice



Employment

Community-Based

- Allow local community to identify poor households
- Idea: local community has much more information than central government
  - This is the premise behind informal insurance, microfinance, etc.
- Problem:
  - If you are using this information to target beneficiaries, this information may not get revealed. Instead, elites may capture the project.
  - Potential tradeoff: better local information vs. more elite capture
- Some existing evidence that communities do know more (Alderman, Galasso and Ravallion)

Current research on targeting

- Alatas, Banerjee, Hanna, Olken, and Tobias (in the field next month!)
- Randomized experiment will compare three targeting methods:
  - Proxy-means test
  - Community ranking
  - Hybrid: community ranking, followed by proxy-means test on bottom 50%.
- Will test corruption in PMT, elite capture of community, and whether hybrid reduces elite capture of ranking process
- To evaluate, we will first conduct household survey to get consumption data, as well as data on family links to village elites and subjective rankings of poverty of other household members
- Stay tuned.

Adding it all up

- Olken (2007) analysis of targeted subsidized rice program in Indonesia
  - In theory, proxy-means test to determine eligibility. Eligible households receive 20kg of subsidized rice per month. Subsidy value about \$4/month, or 9% of HH expenditures for median eligible household.
  - In reality, local officials ignored official criteria and chose beneficiaries.
  - In addition, there was substantial corruption – at least 18% of rice went missing.
- To add this up, calculate social welfare under alternative scenarios:
  - CRRA utility function $u = \frac{c^{1-\rho}}{1-\rho}$
  - Assume all stolen rice goes to richest household in village.
  - Program financed through consumption tax (VAT). Use alternate estimates for marginal cost of public funds (typical developed country estimate: approx 1.3), which measures deadweight loss of taxation
- Normalize social welfare so that complete waste (throw the money in the ocean) = 0% and perfect targeting of transfer = 100%.

Adding it all up

- Local reallocation improved welfare, but corruption may have made program not worthwhile
- Most of the potential gains from redistribution not captured by either PMT or local targeting

Comparing Costs and Benefits

Allocations:		Utilitarian, CRRA utility ρ=1 (% of welfare maximizing utility)	Utilitarian, CRRA utility ρ=2 (% of welfare maximizing utility)
Program	Actual allocation	52.23	35.31
	Actual allocation, no corruption	62.06	42.73
	Official eligibility guidelines	60.90	42.10
No program	Consumption tax, MCF = 1.00	46.90	24.68
	Consumption tax, MCF = 1.20	56.25	29.59
	Consumption tax, MCF = 1.40	65.59	34.48
	Consumption tax, MCF = 1.60	74.91	39.36
Baselines	Pure waste	0.00	0.00
	Welfare maximizing	100.00	100.00

Figure by MIT OpenCourseWare.

Concluding thoughts

- Common theme for taxation and redistribution: lack of information
    - True everywhere, but particularly true in developing countries
    - As a result, tax and redistribution policies look very different
- More broadly, PF and development is a very open area, so lots of room for potential research

Roy's identity details

- Recall

$$V\left(p, y_E\right) = V\left(q, y\right)$$

- Implicit function theorem implies

$$\frac{\partial y_E}{\partial q_i} = \frac{\frac{\partial V(q,y)}{\partial q_i}}{\frac{\partial V(p,y_E)}{\partial y_E}}$$

- Roy's identity implies

$$\frac{\partial y_E}{\partial q_i} = \frac{-\frac{\partial V(q,y)}{\partial y} x_i\left(q, y\right)}{\frac{\partial V(p,y_E)}{\partial y_E}}$$

14.771 Development Economics: Microeconomic Issues and Policy Models
Fall 2008

1 The Family

- A family consists two people F and M with utility functions $U_F(\mathbf{q},\mathbf{a}), U_M(\mathbf{q},\mathbf{a})$, where $\mathbf{q} = (\mathbf{q}_F, \mathbf{q}_M, \mathbf{Q})$ is a vector of amounts of private consumption goods for the two people and the amount of public consumption goods $(\mathbf{Q})$ and $\mathbf{a} = (\mathbf{a}_F, \mathbf{a}_M, \mathbf{A})$ likewise, is a vector of actions that they each can take and a public action.

- Let $\widetilde{\mathbf{p}}= (\mathbf{p}, \mathbf{p}, \mathbf{P})$ be the vector of all the prices of consumption goods. Then the budget constraint for the family:

$$\widetilde{\mathbf{p}}\mathbf{q} = x = x_F(\mathbf{a}) + x_M(\mathbf{a}) + \boldsymbol{\phi}_F + \phi_M$$

  where ϕ_F and ϕ_M are the unearned incomes and $x_F(\mathbf{a})$ and $x_M(\mathbf{a})$ are the earned incomes assigned by the existing system of property rights to F and M. Does not mean M earns the income that is assigned to him.

- Special case 1: $\mathbf{a}$ is **pure investment**. In this case $\partial U_F(\mathbf{q}, \mathbf{a})/\partial \mathbf{a}$ and $\partial U_M(\mathbf{q}, \mathbf{a})/\partial \mathbf{a}$ are both zero. $\mathbf{a}$ only enters through $x_F(\mathbf{a})$ and $x_M(\mathbf{a})$:what is an example?

- Special case 2: **Multiplicative Separability:** $U_F(\mathbf{q}, \mathbf{a}) = U_F(\mathbf{q})\mathbf{g}_F(\mathbf{a})$

1.1 The Unitary Model

- $U_F(\mathbf{q},\mathbf{a}) \equiv U_M(\mathbf{q},\mathbf{a}) = U(\mathbf{q},\mathbf{a})$, i.e identical preferences.

- The consumption decision: Maximize $U(q,\mathbf{a})$ over $\mathbf{q}$ subject to $\widetilde{\mathbf{p}}\mathbf{q} = x_F(\mathbf{a}) + x_M(\mathbf{a}) + \boldsymbol{\phi}_F + \phi_M$. Pareto Optimality by construction.

- FOC

$$\begin{aligned}
\frac{\partial U(\ \mathbf{q},\mathbf{a})}{\partial\ \mathbf{q}} &= \boldsymbol{\lambda}\widetilde{\mathbf{p}} \\
\frac{\partial U(\ \mathbf{q},\mathbf{a})}{\partial \mathbf{a}} &= -\boldsymbol{\lambda}(x'_F(\mathbf{a}) + x'_M(\mathbf{a})) \\
x_F(\mathbf{a}) + x_M(\mathbf{a}) + \boldsymbol{\phi}_F + \phi_M &= \widetilde{\mathbf{p}}\mathbf{q}
\end{aligned}$$

1.2 The fixed bargaining power collective model

- $U_F(\ q, \mathbf{a}) \neq U_M(\ q, \mathbf{a})$, i.e non-identical preferences.

- The family maximizes $U_F(\ q, \mathbf{a}) + \mu U_M(\ q, \mathbf{a})$, where μ is bargaining weight.

- The key assumption is that μ is independent of $x, \mathbf{a}, \mathbf{q}$. One possible scenario is that μ is chosen first, then all the other decisions are taken.

- Given that μ is a constant, the decision taken by the family decision must be Pareto efficient. Why?

- The family's decision: Maximize $U_F(\ q, \mathbf{a}) + \mu U_M(\ q, \mathbf{a})$ over $q, \mathbf{a}$ subject to $\tilde{\mathbf{p}}\ q = x_F(\mathbf{a}) + x_M(\mathbf{a})$. FOC

$$\begin{aligned}
\frac{\partial U_F(\ \mathbf{q}, \mathbf{a})}{\partial\ \mathbf{q}} + \boldsymbol{\mu}\frac{\partial U_M(\ \mathbf{q}, \mathbf{a})}{\partial\ \mathbf{q}} &= \boldsymbol{\lambda}\tilde{\mathbf{p}} \\
\frac{\partial U_F(\ \mathbf{q}, \mathbf{a})}{\partial \mathbf{a}} + \boldsymbol{\mu}\frac{\partial U_M(\ \mathbf{q}, \mathbf{a})}{\partial \mathbf{a}} &= -\boldsymbol{\lambda}(x'_F(\mathbf{a}) + x'_M(\mathbf{a})) \\
x_F(\mathbf{a}) + x_M(\mathbf{a}) + \boldsymbol{\phi}_F + \phi_M &= \tilde{\mathbf{p}}\mathbf{q}
\end{aligned}$$

- Notice that this is formally identical to the unitary model, with

$$U_F(\ \mathbf{q}, \mathbf{a}) + \mu U_M(\ \mathbf{q}, \mathbf{a}) = U_C(\ \mathbf{q}, \mathbf{a})$$

- These two models are not easy to distinguish unless we can observe individual choice behavior outside the usual context of family decision-making (i.e. if you offer choices to one of the members without telling the other).

1.3 Testable Implications of the Model

1.3.1 Income pooling tests

Recall the FOC

$$\begin{aligned}
\frac{\partial U_F(\mathbf{q},\mathbf{a})}{\partial \mathbf{q}} + \boldsymbol{\mu}\frac{\partial U_M(\mathbf{q},\mathbf{a})}{\partial \mathbf{q}} &= \boldsymbol{\lambda}\tilde{\mathbf{p}} \\
\frac{\partial U_F(\mathbf{q},\mathbf{a})}{\partial \mathbf{a}} + \boldsymbol{\mu}\frac{\partial U_M(\mathbf{q},\mathbf{a})}{\partial \mathbf{a}} &= -\boldsymbol{\lambda}(x_F'(\mathbf{a}) + x_M'(\mathbf{a})) \\
x_F(\mathbf{a}) + x_M(\mathbf{a}) + \boldsymbol{\phi}_F + \phi_M &= \tilde{\mathbf{p}}\mathbf{q}
\end{aligned}$$

- Suppose there are different families with the same $x_F(\mathbf{a})$ and $x_M(\mathbf{a})$,and the same total income, but in some of them ϕ_F is large and in others ϕ_M is large.

- Then if they have the same bargaining power, same production technology and same preferences, and face the same prices, they will make the same choices.

- How would we actually test this?

- One problem with this is that we do not usually observe production and utility functions.

- Under what conditions is this not a problem?
  - The windfall test of income pooling

- What do you if windfall shocks are not available?
  1. Strong Separability:

$$\begin{aligned}\partial U_F(\mathbf{q},\mathbf{a})/\partial q_i &= \frac{\partial u_F(\mathbf{q})}{\partial q_i} g(\mathbf{a}),\\ \partial U_M(\mathbf{q},\mathbf{a})/\partial q_i &= \frac{\partial u_M(\mathbf{q})}{\partial q_i} g(\mathbf{a})\end{aligned}$$

     note $g(\mathbf{a})$ does not have F or an M subscript.

Then

$$\begin{aligned} \frac{\partial u_F(\mathbf{q})}{\partial \mathbf{q}} + \boldsymbol{\mu} \frac{\partial u_M(\mathbf{q})}{\partial \mathbf{q}} &= \boldsymbol{\lambda} \frac{\widetilde{\mathbf{p}}}{g(\mathbf{a})} \text{ and} \\ x_F(\mathbf{a}) + x_M(\mathbf{a}) + \boldsymbol{\phi}_F + \phi_M &= \widetilde{\mathbf{p}}\mathbf{q} \end{aligned}$$

can be solved to get $\mathbf{q}(x_F(\mathbf{a}) + x_M(\mathbf{a}) + \boldsymbol{\phi}_F + \phi_M)$.

In this case $\mathbf{q}$ depends only on total family income, $x_F(\mathbf{a}) + x_M(\mathbf{a}) + \boldsymbol{\phi}_F + \phi_M$. *The intra-family distribution of income does not matter*.

What do we need to know to do this?

2. The Ratio Test

Using separability + purely private preferences and

$$\begin{aligned} U_F(\mathbf{q}, \mathbf{a}) &= u_F(\mathbf{q}_F) g_F(\mathbf{a}), \\ U_M(\mathbf{q}, \mathbf{a}) &= u_M(\mathbf{q}_M) g_M(\mathbf{a}). \end{aligned}$$

No spillovers, no family public consumption goods.

In this case by separability the FOC reduces to

$$
\begin{aligned}
g_F(\mathbf{a})\partial U_F(\ \mathbf{q}_F)/\partial\ \mathbf{q}_F &= \boldsymbol{\lambda}\mathbf{p} \\
g_M(\mathbf{a})\boldsymbol{\mu}\partial U_M(\ \mathbf{q}_M)/\partial\ \mathbf{q}_M &= \boldsymbol{\lambda}\mathbf{p} \\
x_F(\mathbf{a})+x_M(\mathbf{a})+\boldsymbol{\phi}_F+\phi_M &= \widetilde{\mathbf{p}}\ \mathbf{q}
\end{aligned}
$$

Therefore $\frac{\partial U_F(\ \mathbf{q}_F)/\partial q_{iF}}{\partial U_F(\ \mathbf{q}_F)/\partial q_{jF}} = \frac{p_i}{p_j} = \frac{\partial U_M(\ \mathbf{q}_M)/\partial q_{iM}}{\partial U_M(\ \mathbf{q}_M)/\partial q_{jM}}$,which implies that the marginal rates of substitution between any two goods is independent of who has bargaining power, as long as there is efficient bargaining. Used for tests of efficiency.

How do we implement this test? Hint: Assume that both U_F and U_M are CRRA. Then derive the "ratio" test.

how robust is this test?

is there a more robust test?

Note that this test works as long as there are a subset of goods for which separability is a reasonable assumption.

3. The investment test

The investment model.

Recall the FOCs in the bargaining model

$$\begin{aligned}
\frac{\partial U_F(\mathbf{q},\mathbf{a})}{\partial \mathbf{q}}+\boldsymbol{\mu}\frac{\partial U_M(\mathbf{q},\mathbf{a})}{\partial \mathbf{q}} &= \boldsymbol{\lambda}\widetilde{\mathbf{p}} \\
\frac{\partial U_F(\mathbf{q},\mathbf{a})}{\partial \mathbf{a}}+\boldsymbol{\mu}\frac{\partial U_M(\mathbf{q},\mathbf{a})}{\partial \mathbf{a}} &= \boldsymbol{\lambda}(x'_F(\mathbf{a})+x'_M(\mathbf{a})) \\
x_F(\mathbf{a})+x_M(\mathbf{a})+\boldsymbol{\phi} &= \widetilde{\mathbf{p}}\,\mathbf{q}
\end{aligned}$$

- In the investment model $\frac{\partial U_F(\mathbf{q},\mathbf{a})}{\partial \mathbf{a}},\frac{\partial U_M(\mathbf{q},\mathbf{a})}{\partial \mathbf{a}}$ are zero. Hence it must be true that

$$x'_F(\mathbf{a})+x'_M(\mathbf{a})=\mathbf{0}$$

This is what Chris Udry, for example, tests.

1.4 The collective model with endogenous bargaining power: Browning-Chiappori

- Gets rid of the assumption that bargaining power is a constant

- $\mu = \mu(x_F(\mathbf{a}), x_M(\mathbf{a}), \boldsymbol{\phi}_F, \phi_M, \mathbf{p})$

- What are properties of the Slutsky Matrix when μ is a constant?

- What is the SR1 property?

- What is the intuition for it?

- What is being tested here?

1.5 What can these models really tell us?

An incomplete contract approach

- What is the alternative to the collective model with endogenous bargaining power?

- One possibility is that the intrafamily contract is not enforceable.

- For a certain good i the husband consumes $\lambda_M^i q^i$ while the wife consumes $(1-\lambda_M^i)q^i$ when the total amount of good i purchased is q^i, irrespective of the relative intensity of their preferences.

- Do we have Pareto Efficiency here?

- Do we have the SR1 property?

- To see that the SR1 property still holds, define a new utility function for each member of the family:

$$\begin{aligned} & U_F(q_F^1, ..q_F^{i-1}, (1-\lambda_M^i)q^i, ..q_F^n, q_M^1, ..q_M^{i-1}, \lambda_M^i q^i, ..q_M^n, \mathbf{Q}, \mathbf{a}) \\ = \; & W_F(q_F^1, ..q_F^{i-1}, q_F^{i+1}, ..q_F^n, q_M^1, ..q_M^{i-1}, \lambda_M^i q^i, ..q_M^n, q^i, \mathbf{Q}, \mathbf{a}) \end{aligned}$$

  and likewise

$$\begin{aligned} & U_M(q_F^1, ..q_F^{i-1}, \lambda_M^i q^i, ..q_F^n, q_M^1, ..q_M^{i-1}, \lambda_M^i q^i, ..q_M^n, \mathbf{Q}, \mathbf{a}) \\ = \; & W_M(q_F^1, ..q_F^{i-1}, q_F^{i+1}, ..q_F^n, q_M^1, ..q_M^{i-1}, \lambda_M^i q^i, ..q_M^n, q^i, \mathbf{Q}, \mathbf{a}) \end{aligned}$$

- W_F and W_M are just two other utility functions with an additional public good q^i.

- So the SR1 property will continue to hold.

- How about the ratio test?

1.5.1 An interesting incomplete contracts model (Maher-Wells)

- One good, investment model: $U_F = u(\ q_F), U_M = u(\ q_M)$.

- The budget constraint is that

$$q_F + q_M = x_F(a_F) + x_M(a_M).$$

- If μ is, as before, fixed, then the family will set $x'_F(a_F) = x'_M(a_M) = 0$ and then distribute consumption.

- Now let μ be dtermined after the investment is made but before consumption is chosen. Let $\mu(\frac{a_F}{a_M})$ and $\mu' < 0$. If the woman invests then she become more powerful. One reason may be that she can just walk off with her $x(a_F)$. i.e her outside option is $u(x(a_F))$ and the bargaining has to give her at least this.

1.5.2 The incomplete contract approach continued

- Maximizing $u(\ q_F) + \mu(a_F/a_M)u(\ q_M)$ subject to the budget constraint yields two functions

$$\begin{aligned} u_F^* &= u_F^*(a_F/a_M, x_F(a_F) + x_M(a_M)) \\ u_M^* &= u_M^*(a_M/a_F, x_F(a_F) + x_M(a_M)) \end{aligned}$$

- Now suppose that F chooses a_F to maximize u_F^* and likewise for M. We assume non-cooperative behavior.

- Since u_F^* is increasing in a_F/a_M and u_M^* is increasing in a_M/a_F, both F and M will over-invest, i.e. $x_F' < 0$ and $x_M' < 0$.

- Maher-Wells give an example of delayed child-bearing.

- How can we distinguish this approach from the complete contract approach with more complex preferences?

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

Testing Household Models

Esther Duflo

14.771

Outline

- Is the household unitary?
- Is the household efficient?
- What next?

Is the Household Unitary?

- Do things other than prices and overall resources ("distribution factors") enter in the production function
- Most tests are test of "income pooling": Does the identity of a transfer recipient matter?
- Other things can influence distribution inside the household:
  - Divorce Laws (Chiappori-Fortin-Lacroix)
  - Marriage markets (Angrist; Lafortune)
  - Labor market
  - Assets brought to the wedding and that spouse retains control of (Thomas-Frankenberg-Contreras)

Testing for income pooling

- Large literature testing for income pooling (Duncan Thomas) You may want to run:

$$z_i = \alpha + \beta y_i^f + \gamma y_i^m + X_i \beta + \epsilon_i$$

  for some outcome z_i, y_i^f is female income, y_i^m male income.
- A number of empirical difficulties with this regression:
  - Joint determination of incomes and consumption: Thomas proposes to use "unearned income" instead.
  - Omitted variables
    - Individual level omitted variables
    - Marriage market: distribution of income reveals something on the spouse

A test of income pooling: Duflo, 2000

- Ideal experiment: an unexpected permanent transfer occurring after marriage (e.g. random allocation of CCT transfers to women or to men: ongoing in Morocco)
- Old Age Pension in South Africa is an approximation:
  - Small extended to Black After the end of Apartheid (1991)
  - Men above 65 and Women above 60 are eligible conditional on a loosely applied means test: 85% of age eligible people are getting it
  - Twice median income per capital in rural areas when it started
  - Many old persons live in 3 generations households, one third of children 0 to 5 lived with a pension recipient in 1993
- Question: Was money spent differently in a household if it was received by a man vs a woman.

Empirical strategy

- Outcome of interest: Children's weight-for-age and height-for-age
- Children who live with pensioners live in different households than those who don't (extended families are poorer, more rural, etc.).
- This may also differ for female vs male.
- Two strategies:
  - "Regression Discontinuity Design" using the age cutoffs for pension recipient for weight-for-height
  - Difference-in-difference for the height-for-age

Weight for height

- Weight for height is flow measure of nutrition, will respond fast to any change in nutrition level
- Idea: Compare children living in 3 generation households with grandmothers eligible vs just a little too young to be eligible; Same thing for grandfathers

$$w_{ijk} = \pi_f E_f + \pi_m E_m + \sum_{j=1}^{4} \gamma_l \mathbf{1}_{(l=k)} + W_{ijk}\lambda + X_{ijk}\delta + \omega_{ijk} \quad (1)$$

- Results
  - Positive impact of grandmother pension on girls, no effect of grandfather's pension
  - Mother's mother pension matters

Height for Age

- Potential problems with Weight for age regressions:
  - Remaining differences between families
  - Endogenous family composition
- Height for age is a stock measure of nutrition, will respond slowly, and no catch up till later of growth deficit in early childhood
- Idea: use the older children as control for younger children in a DD framework: is there a bigger difference between older and younger children in households that are
- Graphs
- Regression:

$$
\begin{aligned}
h_{ijk} \quad = \quad & \pi_f(YOUNG * E_f) + \pi_m(YOUNG * E_m) + \beta_f E_f + \beta_m E_m + \\
& \sum_{l=1}^{4} \gamma_j 1_{(l=k)} + X_{ijk}\delta + \sum_{l=1}^{4} 1_{(l=k)} * X_{ijk}\lambda_j + \epsilon_{ikk} \qquad (2)
\end{aligned}
$$

- Results similar as for weight for height.

Household Efficiency: Ratio tests

- Thomas, Frankenberg, Contreras (2002)
- You have seen the theory beyond these types of test in the previous lecture
- Take two measures of child health, ϕ_k and ϕ'_k, and let A_1 the asset that the wife took to the marriage and A_2 the asset that the husband took to the marriage
- Efficient implies:

$$\frac{\frac{\partial \phi_k}{\partial A_1}}{\frac{\partial \phi_k}{\partial A_2}} = \frac{\frac{\partial \phi_{k'}}{\partial A_1}}{\frac{\partial \phi_{k'}}{\partial A_2}} \tag{3}$$

- Results: Coefficients estimates and Ratio Tests
- No rejection
- Limits of these types of tests:
  - Inherit all the income pooling problems
  - Power (power of overid test to reject is low)

Household Efficiency: Production Efficiency

- Udry (1996)
- Intuition: Separability results: An efficient household should maximize the resources available, and *then* share them.
- Burkina Faso: women and men farm different plot
- Prediction of efficiency: conditioning for the type of crops farmed on each farm, and the productivity of the plot, the yield on women's and men's plots should be the same
- Test this prediction and strongly reject: ▸ Output ▸ Inputs
- This seems to be coming at least in part from much lower use of inputs on women's farm.
- Obvious ways to reconcile with efficiency do not seem to explain the results away

- What is the likely source of violation of efficiency here?
  - Household looks at *income* brought by each household member (rather than potential income). Household member invest to increase their *share* of the income (not only maximize total pie), to influence their bargaining power.
  - Note that this means that husband should buy out the wife (and promise her a utility stream to compensate her).
- Other setting where this "buying out" policy would be efficient: Goldstein-Udry (women are less likely to fallow their land because their property rights are not very secure).

Household Efficiency: Insurance

- Another prediction of a pareto efficient household is that household members should insure each other
- In other words, the pareto weights should not fluctuate with year to year variation in income.
- Women and Men (tend to) grow different crop, on their different farm.
- A special crop is Yam, which is to be used by men for household public goods.
- We can compute proxies for male and female income (and yam income) by aggregating crop income across different crops.
- Haddad and Hodinott run:

$$log(c_{it}) = \alpha + \beta y_{fit} + \gamma y_{mit} + \delta y_{yit} + \epsilon_{it}$$

what are the various reasons why we may expect β and γ to differ?

- I first predict $y_{si2} - y_{si1}$, for s in $\{m, f, y\}$ as a function of rainfall ▸ First Stage and form predicted value of those difference $\hat{\Delta} y_{si} = y_{si2} - y_{si1}$, and I run

$$\Delta(log(c_i)) = \alpha + \beta\hat{\Delta}y_{fi} + \gamma\hat{\Delta}y_{mi} + \delta\hat{\Delta}y_{yi} + \epsilon_i$$

  in a Pareto-efficient model, why would the coefficient β, γ and δ differ?
- What test of Pareto-efficiency does this suggest?
- Consumption of particular goods should change only to the extent that total expenditure changes.
- Two steps:
  1. Run the same regression with total expenditures are the dependent variable

$$\Delta(log(x_i)) = \pi_1 + \pi_2\hat{\Delta}y_{fi} + \pi_3\hat{\Delta}y_{mi} + \pi_4\hat{\Delta}y_{yi} + \epsilon_i$$

  2. calculate the ratios: $\frac{\beta}{\pi_1}$, $\frac{\gamma}{\pi_2}$, $\frac{\delta}{\pi_3}$. They should all be equal.

Results Interpretation

- Results Rejection of equality of ratio
- Does not seem to be explained by obvious failure of identification
- Is this a labor market failure?
- Can this be due to lack of observability of the output?
- Can this be due to moral Hazard?
- Why do household keep separate mental account?
  - Incomplete contracting in the household: constant negotiations of what transfers should be in a given period are very difficult.
  - Households members decide instead of very simple rules they follow, and would be subject to strong punishment if they re-negociated upon. This allows for insurance against mis-behavior (and perhaps avoids the unpleasantness of negociating).

Duflo (2003)

Effect of the Old Age Pension Program on Weight for Height: Ordinary Least Squares and Two-Stage Least Squares Regressions

Variable	*Ordinary least squares*						*Two-stage least squares*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Girls							
Eligible household	0.14 (0.12)	0.35* (0.17)	0.34* (0.17)				
Woman eligible				0.24* (0.12)	0.61* (0.19)	0.61* (0.19)	1.19* (0.41)
Man eligible				-0.011 (0.22)	0.11 (0.28)	0.056 (0.19)	-0.097 (0.74)
Observations	1,574	1,574	1,533	1,574	1,574	1,533	1,533
Boys							
Eligible household	0.0012 (0.13)	0.022 (0.22)	0.030 (0.24)				
Woman eligible				0.066 (0.14)	0.28 (0.28)	0.31 (0.28)	0.58 (0.53)
Man eligible				-0.059 (0.22)	-0.25 (0.34)	-0.25 (0.35)	-0.69 (0.91)
Observations	1,670	1,670	1,627	1,670	1,670	1,627	1,627
Control variables							
Presence of older members	No	Yes	Yes	No	Yes	Yes	Yes
Family background variables	No	No	Yes	No	No	Yes	Yes
Child age dummy variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes

** Significant at the five percent level*

Note: The instruments in column 7 are woman eligible and man eligible. Standard errors (robust to correlation of residuals within house holds and heteroskedasticity) are in parentheses.

Figure by MIT OpenCourseWare.

Duflo (2003)

Effect of Pension Eligibility on Weight for Height by Gender of the Intermediate Generation: Ordinary Least Squares Regressions

Variable	*Girls*	*Boys*
Mother's mother Eligible	0.48* (0.21)	0.099 (0.27)
Father's mother Eligible	0.15 (0.25)	0.29 (0.30)
Mother's father Eligible	0.097 (0.34)	0.00052 (0.43)
Father's father Eligible	0.22 (0.48)	0.25 (0.44)
Observations	1,457	1,552
Control variables		
Presence of older members	Yes	Yes
Family background variables	Yes	Yes
Age dummy variables	Yes	Yes

** Significant at the five percent level.*

Note: Standard errors (robust to correlation of residuals within households and heteroskedasticity) are in parentheses.

Figure by MIT OpenCourseWare.

	Treatment variable			
	Eligibility	*Eligibility*	*Old grandparent*	*Receives pension*
	Ordinary least squares			*Two-stage least squares*
	(1)	*(2)*	*(3)*	*(4)*
Girls				
Eligible household * YOUNG	0.68* (0.37)			
Woman treatment variable * YOUNG		0.71* (0.34)	0.40 (0.27)	1.16* (0.56)
Man treatment variable * YOUNG		0.097 (0.57)	-0.12 (0.35)	-0.071 (0.95)
Eligible household	-0.17 (0.16)			
Woman pension variable		-0.15 (0.17)	-0.039 (0.13)	-0.15 (0.17)
Man pension variable		-0.11 (0.24)	0.027 (0.15)	-0.11 (0.24)
Observations	1,533	1,533	1,533	1,533

Figure by MIT OpenCourseWare.

Duflo (2003)

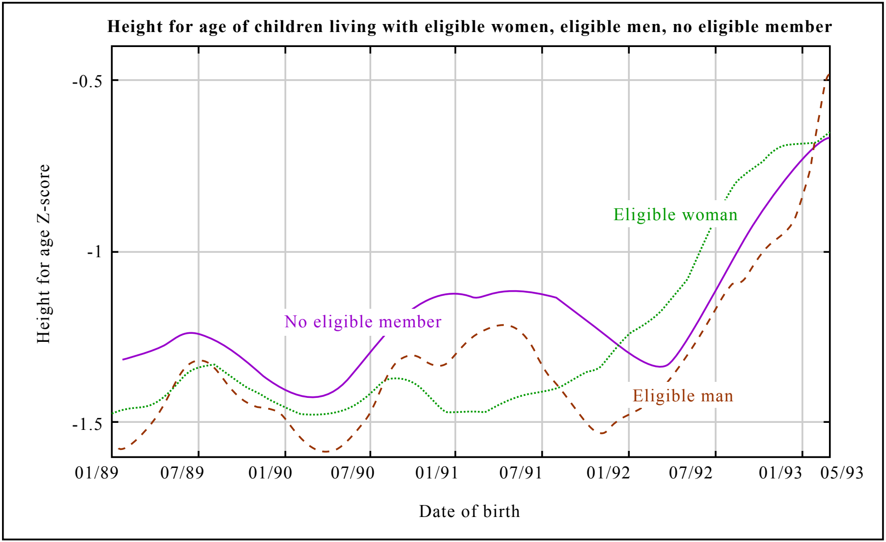


Figure by MIT OpenCourseWare.

Thomas, Frankenberg, and Conteras (2002)

Impact of Parental Assets at Marriage on Child Morbidity:
OLS and Fixed Effects Estimates (*100)
Java and Sumatra

	Sons	*Daughters*	*Difference*	
			OLS	*Fixed effects*
Cough				
Paternal assets at marriage	0.135 [2.60]	0.011 [0.14]	0.124 [1.30]	0.119 [1.37]
Maternal assets at marriage	-0.093 [1.09]	0.143 [1.53]	-0.237 [1.86]	-0.236 [2.78]
χ^2(asset effects=0)	3.90 [0.02]	1.21 [0.30]	2.42 [0.09]	4.73 [0.01]
χ^2(asset effects equal)	5.08 [0.02]	1.04 [0.31]	4.82 [0.03]	8.36 [0.00]
F (all covariates)	10.46 [0.00]	2.60 [0.00]	7.10 [0.00]	2.78 [0.00]
R^2	0.096	0.085	0.091	0.686
Fever				
Paternal assets at marriage	0.068 [0.74]	0.075 [0.90]	-0.007 [0.05]	-0.026 [0.25]
Maternal assets at marriage	0.029 [0.33]	0.224 [2.44]	-0.195 [1.53]	-0.186 [2.48]
χ^2(asset effects=0)	0.36 [0.70]	3.67 [0.03]	1.20 [0.30]	3.21 [0.04]
χ^2(asset effects equal)	0.09 [0.77]	1.29 [0.26]	1.01 [0.32]	1.46 [0.23]
F (all covariates)	5.50 [0.00]	3.01 [0.00]	4.50 [0.00]	2.53 [0.00]
R^2	0.080	0.083	0.082	0.655

Notes: Sample size: 601 sibling pairs. Standard errors below coefficient estimates; p-values below test statistics. Variance-covariances matrices computed by method of infinitesimal jackknife.

Thomas, Frankenberg, and Conteras (2002)

Impact of Parental Assets at Marriage on Child Morbidity:
OLS and Fixed Effects Estimates (*100)
Java and Sumatra

	Sons	*Daughters*	*Difference*	
			OLS	*Fixed effects*
Diarrhea				
Paternal assets at marriage	-0.002 [0.03]	0.072 [0.85]	-0.074 [0.69]	-0.079 [1.39]
Maternal assets at marriage	-0.042 [1.13]	-0.018 [0.45]	-0.024 [0.43]	-0.017 [0.42]
χ^2(asset effects=0)	0.64 [0.53]	0.45 [0.64]	0.320 [0.73]	0.980 [0.38]
χ^2(asset effects equal)	0.29 [0.59]	0.89 [0.35]	0.170 [0.68]	0.970 [0.33]
F (all covariates)	2.59 [0.00]	1.87 [0.01]	2.180 [0.00]	2.030 [0.00]
R^2	0.071	0.062	0.067	0.682
Other				
Paternal assets at marriage	0.066 [1.05]	0.096 [1.19]	-0.030 [0.30]	-0.063 [0.61]
Maternal assets at marriage	0.066 [1.24]	-0.023 [0.31]	0.089 [0.97]	0.110 [1.57]
χ^2(asset effects=0)	1.31 [0.27]	0.73 [0.48]	0.500 [0.61]	1.340 [0.26]
χ^2(asset effects equal)	0.00 [1.00]	1.08 [0.30]	0.720 [0.40]	1.750 [0.19]
F (all covariates)	6.80 [0.00]	2.52 [0.00]	4.570 [0.00]	1.910 [0.00]
R^2	0.081	0.044	0.064	0.684

Notes: Sample size: 601 sibling pairs. Standard errors below coefficient estimates; p-values below test statistics. Variance-covariances matrices computed by method of infinitesimal jackknife.

Figure by MIT OpenCourseWare.

Thomas, Frankenberg, and Conteras (2002)

Tests for Pareto Efficiency in Household Allocations
Java and Sumatra
Ratio of Effects of Paternal to Maternal Assets at
Marriage and Non Linear Wald Tests for Equality of Ratios

Ratios of asset effects: α_1/α_2	
Cough	-0.50
Fever	0.14
Diarrhea	4.65
Other	-0.57

Pair-wise tests for equality of ratios: $\chi^2{}_1$

	Fever	*Diarrhea*	*Other*
Cough	0.90 [0.14]	1.44 [0.23]	0.00 [0.96]
Fever		1.43 [0.23]	0.41 [0.52]
Diarrhea			1.10 [0.29]

Joint tests for equality of all ratios: $\chi^2{}_5$

	2.52 [0.77]	

Notes: P-values below test statistics. Variance-covariances matrices computed by method of infinitesimal jackknife.

Figure by MIT OpenCourseWare.

OLS Fixed-Effect Estimates of the Determinants of Plot Yield and Ln(Plot Output) (x 1,000 FCFA)
Dependent Variable: Value of Plot Output/Hectare

	Household-year-crop effects: all crops		*Household-year effects*				*Household-crop-year effects*			
			Millet only		*White sorghum*		*Vegetables*		*All crops: CES**	
	(1)		*(2)*		*(3)*		*(4)*		*(5)*	
Mean of dependent variable	89		31		41		134		1.67	
Gender: (1 = female)	-27.70	(-4.61)	-10.36	(-2.53)	-19.38	(-4.43)	-34.27	(-2.21)	-.20	(-3.56)
Plot size:										
1st decile	133.99	(3.50)	-28.35	(-2.67)	-17.90	(-1.92)	237.10	(4.66)		
2d decile	69.10	(4.38)	8.64	(.82)	52.30	(3.16)	63.97	(2.38)		
3d decile	63.45	(5.52)	16.95	(1.81)	47.68	(4.77)	35.87	(1.52)		
4th decile	34.08	(2.88)	9.79	(1.12)	26.73	(3.12)	4.21	(.18)		
6th decile	-2.04	(-.29)	-.99	(-.11)	-6.38	(-1.16)	-6.65	(-.26)		
7th decile	-13.44	(-1.78)	-13.01	(-1.73)	-11.31	(-1.69)	-33.54	(-.90)		
8th decile	-17.23	(-2.59)	-12.97	(-1.34)	-28.58	(-4.82)	31.04	(.73)		
9th decile	-26.68	(-3.81)	-21.50	(-2.65)	-28.65	(-4.98)				
10th decile	-31.52	(-4.49)	-20.56	(-2.55)	-37.70	(-6.03)				
Ln(area)									.78	(29.52)
Toposequence:										
Uppermost	-41.35	(-2.18)	2.50	(.24)	-14.60	(-1.73)	-131.34	(-1.82)	-.46	(-2.71)
Top of slope	-26.35	(-1.27)	9.53	(.96)	-11.27	(-1.47)	-121.05	(-1.85)	-.29	(-1.92)
Mid-slope	-24.38	(-1.19)	5.39	(.64)	-8.62	(-1.15)	-119.68	(-1.88)	-.28	(-1.97)
Near bottom	-21.70	(-.90)	4.48	(.40)	-5.36	(-.71)	-93.96	(-1.30)	-.18	(-1.27)

Figure by MIT OpenCourseWare.

Udry (1996)

Least-Squares Tobit Fixed-Effect Estimates of the Determinants of Plot Input Intensities

	Household-year-crop effects									
	Male labor per hectare (1)		*Female labor per hectare (2)*		*Child labor per hectare (3)*		*Nonhousehold labor per hectare (4)*		*Manure (1,000) kg per hectare (5)*	
Gender: (1 = female)	-668.47	(-9.60)	70.23	(1.53)	-195.46	(-2.34)	-428.41	(-1.70)	-16.33	(-2.54)
Plot size:										
1st decile	1,209.72	(2.53)	1,462.21	(5.71)	740.80	(1.17)	193.35	(.43)	24.79	(2.42)
2d decile	417.18	(3.25)	1,131.01	(5.82)	143.12	(1.11)	487.39	(1.28)	7.99	(.96)
3d decile	245.94	(2.74)	799.12	(6.72)	133.16	(1.53)	689.39	(1.27)	2.58	(.48)
4th decile	96.53	(1.71)	407.87	(5.02)	72.51	(.68)	378.18	(1.07)	-6.18	(-1.12)
6th decile	-.55	(-.01)	-69.25	(-1.36)	-72.15	(-.98)	57.48	(.80)	-2.14	(-.33)
7th decile	-153.12	(-2.97)	-306.51	(-5.96)	-59.53	(-.60)	65.51	(.64)	-11.08	(-1.54)
8th decile	-375.53	(-6.23)	-386.78	(-6.61)	-184.61	(-1.61)	-43.81	(-.30)	-11.01	(-1.61)
9th decile	-413.36	(-6.79)	-373.57	(-5.16)	-269.99	(-1.83)	-255.15	(-.87)	-11.64	(-1.80)
10th decile	-490.11	(-7.72)	-418.06	(-6.08)	-219.27	(-1.86)	-220.64	(-1.07)	-16.41	(-2.45)
Toposequence:										
Uppermost	41.62	(.35)	-1.92	(-.02)	-55.52	(-.51)	20.20	(.12)	-9.22	(-.62)
Top of slope	29.36	(.30)	91.02	(1.07)	35.15	(.38)	144.02	(.83)	.26	(.02)
Mid-slope	36.08	(.38)	.57	(.01)	.10	(.00)	-15.45	(-.11)	1.14	(.11)
Near bottom	16.42	(.18)	75.94	(.86)	-98.03	(-1.05)	23.27	(.17)	2.88	(-.27)
Soil types:										
3	103.49	(.60)	-31.68	(-.23)	235.74	(.86)	175.29	(.50)	-11.80	(-1.18)
7	-65.79	(-.85)	-30.39	(-.28)	21.88	(.44)	66.04	(.47)	-.07	(-.01)
11	-28.77	(-.09)	-52.06	(-.34)	-778.86	(-4.36)	262.71	(.70)	-.70	(-.08)
12	1,051.98	(.82)	367.34	(1.63)	62.36	(.44)	368.47	(1.13)	16.32	(1.48)
13	274.48	(1.33)	-38.50	(-.29)			-187.07	(-.89)		
21	196.37	(.95)	-53.41	(-.49)	-42.87	(.35)	37.73	(.27)	2.86	(.18)
31	83.16	(1.59)	68.24	(.92)	205.90	(2.29)	115.56	(1.00)	6.43	(1.29)
32	24.77	(.50)	-10.36	(-.15)	173.14	(1.07)	51.08	(-.44)	.73	(.12)
33	250.40	(2.57)	163.76	(1.36)	206.68	(.78)	-113.72	(-.37)	17.28	(1.61)
35	179.46	(1.50)	303.86	(1.90)	248.38	(2.60)	195.14	(.58)	-12.75	(-.94)
37	82.49	(.70)	50.84	(.30)	114.53	(1.19)	31.14	(.20)	8.34	(1.44)
45	78.13	(1.34)	-8.33	(-.10)	79.85	(1.02)	41.90	(.25)	8.00	(1.83)
46	-187.14	(-1.84)	141.73	(.76)	42.70	(.09)	223.23	(1.27)	-15.45	(-.79)
51	95.73	(1.83)	-27.01	(-.33)	2.93	(.05)	126.70	(1.05)	.80	(.17)
Location:										
Compound	35.35	(.78)	37.16	(.90)	-18.82	(-.31)	-162.88	(-1.38)	.99	(.24)
Village	19.69	(.70)	12.18	(.45)	42.92	(.93)	25.80	(.30)	5.86	(1.60)
Mean of dependent variable	427.39		466.18		85.55		84.88		1.70	
When > 0	506.62		517.17		202.88		213.11		7.78	

Note.- This is the least-squares implementation of Honoré's (1992) fixed-effect Tobit estimator. t-ratios are in parentheses.

Figure by MIT OpenCourseWare.

Table 2: First stage summary statistics

	Dependent variables		
	Male cash crop	Current Yam income	Female Income
	(1)	(2)	(3)
F statistics (p value)			
All rainfall variables	1.99	3.50	2.53
are significant	(0.014)	(0.000)	(0.000)
Current year rainfall variables	1.18	3.38	2.43
significant	(0.315)	(0.000)	(0.005)
Past year rainfall variables	2.79	4.64	2.64
significant	(0.005)	(0.000)	(0.001)
Rainfall variables significantly different from:			
Male cash crop	NA		
	2.10		
Yam income	(0.010)	NA	
Female income	2.10	2.38	NA
	(0.009)	(0.002)	

Table 4: Restricted overidentification tests

	Dependent variable: Change in log (item consumption)											
	Total expenditure	Food consumption	Adult goods	Clothing	Prestige goods	Education	Staples	Meat	Vegetables	Processed foods	Purchased foods	Food consumed at home
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
PANEL A												
OLS coefficients:												
Predicted change in male non-yam	0.126	0.062	0.870	-0.164	0.683	-0.101	0.113	0.002	0.345	0.004	-0.029	0.098
income	(0.049)	(0.054)	(0.425)	(0.334)	(0.209)	(0.128)	(0.072)	(0.126)	(0.210)	(0.139)	(0.078)	(0.119)
Predicted change in yam	0.207	0.227	-0.473	0.296	-0.272	0.320	0.345	0.135	0.023	0.122	0.087	0.444
income	(0.037)	(0.041)	(0.320)	(0.252)	(0.158)	(0.108)	(0.054)	(0.096)	(0.159)	(0.105)	(0.059)	(0.090)
Predicted change in female	0.309	0.235	1.537	0.535	0.993	-0.098	0.193	0.492	0.995	0.474	0.412	0.313
income	(0.056)	(0.061)	(0.490)	(0.382)	(0.239)	(0.159)	(0.082)	(0.144)	(0.239)	(0.159)	(0.089)	(0.136)
F tests (p value) :		0.934	5.064	0.514	7.595	2.260	5.870	1.824	3.277	1.397	4.777	1.912
Overidentification		(0.393)	(0.007)	(0.598)	(0.001)	(0.106)	(0.003)	(0.162)	(0.038)	(0.248)	(0.009)	(0.148)
Restriction test												
PANEL B: LAGGED RAINFALL												
OLS coefficients:												
Predicted change in lagged male	0.073	0.039	0.350	0.044	0.047	0.091	0.038	0.150	0.039	0.115	0.155	-0.007
non-yam income	(0.020)	(0.022)	(0.169)	(0.133)	(0.082)	(0.056)	(0.029)	(0.050)	(0.083)	(0.055)	(0.031)	(0.047)
Predicted change in lagged yam	-0.003	0.004	0.008	-0.125	-0.076	-0.031	-0.021	0.015	0.011	0.027	0.024	-0.018
income	(0.009)	(0.009)	(0.073)	(0.059)	(0.036)	(0.029)	(0.013)	(0.022)	(0.036)	(0.024)	(0.013)	(0.021)
Predicted change in lagged female	-0.001	0.018	-0.024	-0.251	-0.289	0.093	0.044	0.023	-0.054	-0.010	0.062	-0.035
income	(0.026)	(0.028)	(0.220)	(0.173)	(0.107)	(0.079)	(0.038)	(0.064)	(0.107)	(0.071)	(0.040)	(0.061)
F tests (p value) :		0.105	0.128	0.254	0.043	0.016	0.049	0.052	0.024	0.058	0.054	0.057
Overidentification		(0.900)	(0.880)	(0.776)	(0.958)	(0.984)	(0.952)	(0.949)	(0.976)	(0.943)	(0.948)	(0.945)

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

14.771: Firms and Contracts Lecture 1

Ben Olken

October 2008

Outline

- Topics I will talk briefly about:
  - Contracts and reputation (today)
  - Implications for corporate finance (next time)

Background

- Basic problem: bad legal system means that contracts are hard to enforce
  - E.g., recovering debts, enforcing contract disputes, etc.
- Of course, this is true everywhere to some degree
  - But this is often thought to be worse in developing countries.
  - See Doing Business 2008
  - (although note that their academic papers find only lukewarm support for this result)

Background

Where is Enforcing Contracts Easy - and Where not?

Easiest	*Rank*	*Most difficult*	*Rank*
Hong Kong, China	1	Central African Republic	169
Luxembourg	2	Belize	170
Latvia	3	Syria	171
Iceland	4	Cameroon	172
Singapore	5	Congo, Dem. Rep.	173
Austria	6	Suriname	174
Finland	7	Bangladesh	175
United States	8	Angola	176
Norway	9	India	177
Korea	10	Timor-Leste	178

Note: Rankings are the average of the country rankings on the procedures, time and cost to resolve a commercial dispute through the courts. See data notes for details.
Source: Doing Business *database.*

Figure by MIT OpenCourseWare.

Background

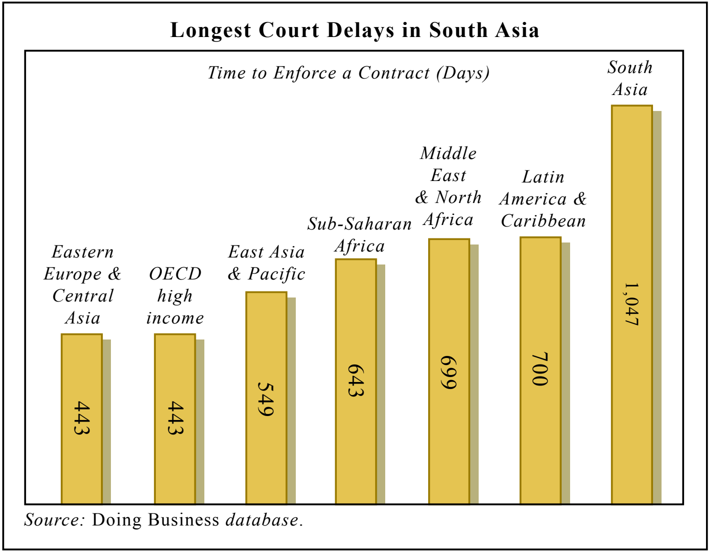


Figure by MIT OpenCourseWare.

Background

Cost (% of claim)

Least		*Most*	
Bhutan	0.1	Comoros	89.4
Iceland	6.1	Cambodia	102.7
China	8.8	Burkina Faso	107.4
Luxembourg	8.8	Papua New Guinea	110.3
United States	9.4	Indonesia	122.7
Norway	9.9	Malawi	142.4
Poland	10.0	Mozambique	142.5
Korea	10.3	Sierra Leone	149.5
Finland	10.4	Congo, Dem. Rep.	151.8
Germany	11.8	Timor-Leste	163.2

Source: Doing Business *database*.

Figure by MIT OpenCourseWare.

Models of Reputations

- So how do we enforce contracts without courts?
- We'll explore several options:
  - Repeated interactions
  - Collective reputations
  - Networks

Tirole (1996)

- Suppose there is a buyer who wants one of two types of goods: red or white
- Quality can be good or bad; this is observable but non-verifiable (i.e. not enforceable by court)
- Payoffs to buyer:
  - If buyer orders red: utility from good quality is H and utility from bad quality is D
  - If buyer orders white: utility from good quality is h and uility from bad quality is d
- Assume $H > h > d > D$. Assume $d \geq 0$.
- Assume that the buyer proposes the equilibrium.

Suppliers

- Suppliers have a cost G per period of supplying the good quality and a cost of 0 of supplying the bad quality.
- Assume the efficient outcome is to produce high quality red, i.e. $H - G > d > h - G$.
- The supplier's outside option is getting zero for ever. The supplier cannot be paid a negative price.
- The relation goes on till the supplier dies, which happens with probability λ each period. No other discounting.

Equilibrium with single buyer and seller

- One-shot game:
  - Supplier always chooses to deviate and produce low-quality
  - Therefore, buyers always order white at $p = 0$.
- Repeated game:
  - Folk-theorem logic: If supplier ever deviates and supplies low quality, order white at $p = 0$ forever.
  - This punishment threat can sustain good behavior.
  - This will be the case if

$$\sum_{n=0}^{\infty} (1-\lambda)^n (p_n - G) \geq p_0$$

  - With constant prices, this is just

$$\frac{p - G}{\lambda} \geq p$$

  - Since seller is willing to pay up to H, this equilibrium exists as long as

$$\lambda < 1 - \frac{G}{H}$$

Equilibrium with multiple types of sellers

- Suppose now there are three types of sellers:
  - Honest (fraction α). Always produce high quality.
  - Dishonest (β). Always produce low quality.
  - Strategic (γ). Do what is in their best interest.
- How does this change the equilibrium?
  - Buyer orders red and starts with $p = 0$. This screens out the dishonest sellers.
  - After this, same equilibrium as before (order white at $p = 0$ if ever get low quality), and order red if continue to get high quality
  - New sellers initially take a loss, but are paid higher prices later to compensate.
  - A buyer who has an established seller may refuse a new supplier even if the price is zero. Why?

Equilibrium with multiple sellers

- Now suppose there are many sellers, and a buyer is matched with a new seller each period.
- If a seller has supplied low quality at least once in the past, buyer finds out with probability x.
- In this model, once a strategic seller supplies low quality he will do so in the future, since his reputation is already tainted.
- Finally assume that the price of red is fixed at B and that of whites is fixed at $b < B$.(In Tirole these are private benefits).

Equilibrium with multiple sellers

- Now there are multiple steady states
- Good equilibrium:
  - Suppose the seller orders red from any buyer with whom he is matched and who is not known to have delivered low quality in the past.
  - Then an untainted strategic seller may produce high quality red if x is high enough and the gap $B - b$ is large enough.
  - Knowing this the buyer will order red as long as there are enough untainted strategic sellers.
- Bad equilibrium:
  - Suppose sellers are expected to always demand white in the future.
  - Then all strategic agents will produce low quality today, since there is no return to preserving reputation.
  - Given this, sellers are better off demanding white today.

Implications

- Persistence:
  - Suppose there is a one-time shock and everyone's reputation is tainted. Buyers and sellers know this.
  - Now the good equilibrium can go away – even for newly born, untainted sellers.
  - Why? Suppose you don't receive a signal that the person is tainted. What is your inference that the person is actually tainted? If λ is sufficiently low, the person is most likely tainted and will deliver low quality.
  - Key intuition: if collective reputation is bad, new untainted people cannot distinguish themselves.
- Information:
  - Information structure (x) is crucial for this model.
  - If x is very high, we are always in good state, since new agents now have an incentive to maintain individual reputation.
  - if x is very low, we are always in bad equilibrium, because we cannot sustain any Folk theorem equilibrium.

Empirical implications

- Sellers may have to take a loss up front in order to establish their relationship. Contracts will therefore change as individual relationships get established.
- Reputations are valuable, and temporary shocks can have long-lasting implications (think of a financial crisis).
- Commercial networks may form where information is more observable (i.e., x is higher).
- Networks also can enhance enforcement by increasing the penalty from default (Kandori 1992, Kranton and Meinhart 2001, and others)

Banerjee and Duflo (2000)

- Setting:
  - Study of the Indian software industry, which produces customized software for large corporations.
  - Software is customized and takes time to produce. The problem is that you don't know how difficult a software project is until you start working on it.
- Firms:
  - As in the model above, there is heterogeneity in the type of firm. Bad firms are inclined to cost overruns.
- Contracting:
  - Contracts are inadequate protection because both sides can claim that the other side was to blame for delays.

Contracts

- There are several two ways to deal with cost overruns
  - Not buying in the future, as in the above model
  - Forcing the firm to pay for it by making it responsible for the overrun. This can be achieved by fixed price contract instead of a time and material contract.
- However a fixed price contract forces the firm to bear all the risk and gives the buyer incentives to misbehave.
- Therefore firms will prefer to move to a time and material contract, but the buyers will not agree unless the firm has a reputation for being good.

Predictions

- Firms that are in a repeat contract is more likely to have time and material contracts.
- Firms that work for 'parent companies' are more likely to have time and material contracts.
- Assume that a firm that does not get some repeat buyers goes out of business. Older firms are therefore less likely to be bad firms. Then older firms are more likely to have a time and material contract.
  - Alternatively, could get a similar result if firms' past behavior with other clients is imperfectly observable to new clients.

Empirics

- Collected data on contracts from a survey of Indian software firms
- Define contract type C_{ic} to be 1 for fixed-price, 2 for mixed, and 3 for time and materials
- Estimate ordered probit

$$C_{ic} = \alpha R_{ic} + \beta X_{ic} + \gamma Z_{ic} + \delta M_{ic} + v_i + \omega_{ic}$$

  where R is reputation variables, X is project characteristics, and M is client characteristics.

- Estimate analogous models for:
  - Whether firm paid for any actual overrun
  - Whether there was an overrun

Results

	Choice of contract ordered probit		*Share of overrun paid by the firm*					
			Unconditional				*Conditional*	
			Random effect		*Fixed effect*		*Random effect*	*Fixed effect*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Reputation								
Young firm	-0.69* (0.25)		15* (8.5)				9.0 (8.6)	
Repeated contract	0.22 (0.24)		-17* (8.8)		-20 (16)		-15* (8.7)	-19 (17)
ISO-certified firm	-0.27 (0.32)		17 (13)				16 (13)	
Internal project		0.87* (0.31)		-25* (11)		-64* (26)		

Figure by MIT OpenCourseWare.

Results

	Total overrun		*Overrun due to the firm*	
	Unconditional	*Conditional*	*Unconditional*	*Conditional*
	Random effect	*Random effect*	*Random effect*	*Random effect*
	(1)	*(2)*	*(3)*	*(4)*
Reputation				
Young firm	-0.48 (5.0)	-3.8 (5.0)	2.5 (3.4)	1.5 (3.5)
Repeated contract	1.8 (4.9)	1.5 (4.8)	-0.92 (3.5)	-1.2 (3.5)
ISO-certified firm	15 (7.9)	16 (7.7)	5.4 (5.4)	6.1 (5.5)

Figure by MIT OpenCourseWare.

Networks

- Greif (1993)
  - Studies the Maghribi traders, a network of Jewish traders
  - Because they shared a common language and were within a common network, it was easier to share information about counter-parties (i.e., high x)
  - If someone deviated, entire network would punish the deviant trader. This created stronger incentives for honest behavior
  - People who were suspected of cheating would have to invest in rebuilding their reputation

Networks

- McMillan and Woodruff (1999)
  - Study provision of trade credit in Vietnam
  - Trade credit requires trust, because you are paid after delivery of goods
  - Networks provide both information and enforcement, as in Greif
- Data on firms in Vietnam
- Key dependent variable: percent of bill paid by customer after delivery
- Key independent variable: talk to other suppliers of customer at least monthly, so in a network of information about customer
- Also examine duration (as in above model)

Results

Manufacturer information:				
Duration of relationship (years)	0.08 (2.96)	0.07 (2.61)	0.07 (2.51)	0.07 (2.42)
Duration^2	-0.005 (2.15)	-0.004 (1.95)	-0.004 (1.74)	-0.004 (1.78)
Visited customer before first sale		0.08 (1.63)	0.07 (1.71)	0.06 (1.33)
Currently visit customer at least weekly		-0.03 (0.46)	-0.06 (1.03)	-0.05 (0.84)
Network membership:				
First information from other manufacturers	0.20 (3.36)	0.16 (2.83)	0.10 (1.99)	0.17 (2.98)
Talk to other suppliers of customer at least monthly		0.19 (2.36)	0.19 (2.63)	0.18 (2.31)
First information from family member	0.04 (0.60)	-0.01 (0.17)	-0.08 (1.34)	-0.13 (2.11)

Figure by MIT OpenCourseWare.

Brands

- All of this has been about reputations vis-a-vis other firms in business to business transactions
- Similar logic may apply to reputations vis-a-vis consumers:
  - Companies invest in building a brand (e.g., "Tata" in India), which is difficult to do
  - And then use that brand to build a wide variety of products
- This provides one potential explanation for why we observe large, diversified conglomerates in developing countries
- More reasons for the presence of conglomerates next time.

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

14.771: Firms and Contracts Lecture 2

Ben Olken

October 2008

Overview

- Last lecture: problems in contract enforcement lead to other types of contract enforcement mechanisms (e.g., reputations, networks)
- This lecture: what are the implications of weak contract enforcement for how firms are structured?
  - Business groups.
    - Some problems with business groups (tunneling)
  - Family firms

Business Groups

- Weak contract enforcement suggests that more is likely to be done within the firm

Business Groups

- Weak contract enforcement suggests that more is likely to be done within the firm
- Other theories might also suggest integration across industries (i.e., unrelated production functions)

Business Groups

- Weak contract enforcement suggests that more is likely to be done within the firm
- Other theories might also suggest integration across industries (i.e., unrelated production functions)
    - Access to finance also means more may be done within the firm in places where finance is less developed (Rajan and Zingales 1998)
    - Branding/reputations (discussed at the end of last lecture) suggests reasons for firms to integrate across sectors

A diversified business group

Diagram of the Slim Helu Group removed due to copyright restrictions.
See Perkins, Morck, and Yeung (2006).

Conglomerate Size and Financial Development

- Cross-country evidence mixed

Conglomerate Size and Financial Development

- Cross-country evidence mixed
  - Note that one shouldn't necessarily count cross-country evidence too much! One should view this as motivation for better micro studies.

Conglomerate Size and Financial Development

- Cross-country evidence mixed
  - Note that one shouldn't necessarily count cross-country evidence too much! One should view this as motivation for better micro studies.
- Acemoglu, Johnson, and Mitton (2005) study vertical integration worldwide
  - For each industry, they use US input-output tables to determine how much input from each industry is required to produce a given type of output
  - They then calculate using each firm's SIC codes what percent of the firm's inputs are produced by industries in which the firm operates

Conglomerate Size and Financial Development

- Cross-country evidence mixed
  - Note that one shouldn't necessarily count cross-country evidence too much! One should view this as motivation for better micro studies.
- Acemoglu, Johnson, and Mitton (2005) study vertical integration worldwide
  - For each industry, they use US input-output tables to determine how much input from each industry is required to produce a given type of output
  - They then calculate using each firm's SIC codes what percent of the firm's inputs are produced by industries in which the firm operates
- Findings:
  - Vertical integration is greater in poorer countries, and in countries with greater cost of contract enforcement
  - Also greater in countries with greater entry cost
  - However, this is due almost entirely to industrial composition
- So it's not clear whether other factors cause these industries to be more appropriate for developing countries, or vice-versa

Acemoglu, Johnson, and Mitton Results

- Actual vertical integration

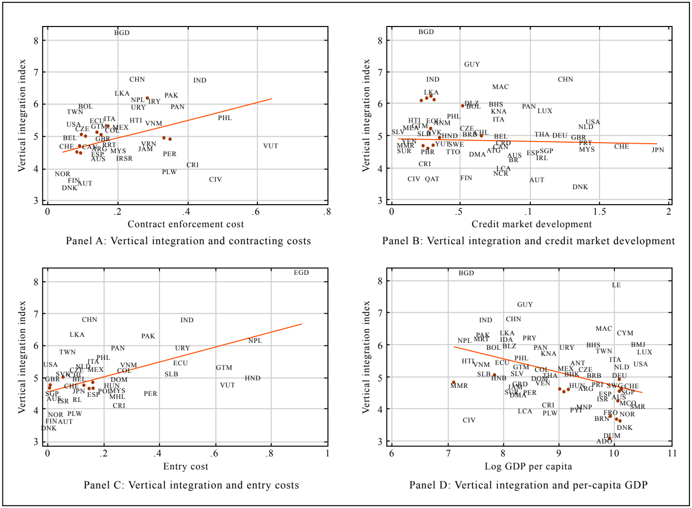
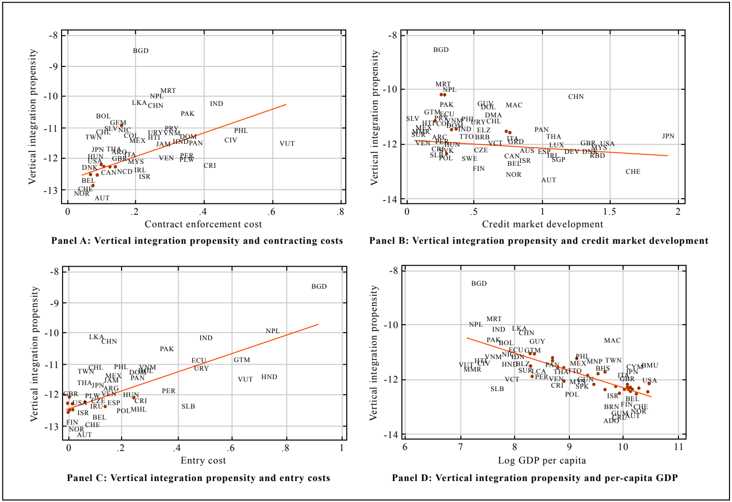


Figure by MIT OpenCourseWare.

- Vertical integration predicted by industry mix



Panel A: Vertical integration propensity and contracting costs

Panel B: Vertical integration propensity and credit market development

Panel C: Vertical integration propensity and entry costs

Panel D: Vertical integration propensity and per-capita GDP

Figure by MIT OpenCourseWare.

Khwaja, Mian and Qamar (2008)

- What are the benefits of being in a network?

Khwaja, Mian and Qamar (2008)

- What are the benefits of being in a network?
- Data on networks:
    - Authors have data on directors of all Pakistani companies, public and private
    - Define two firms as connected if they share a common director
    - Define two firms are in the same network if they can be linked through connected firms
    - Since they have the entire universe of firms, they can construct networks for the whole economy

Khwaja, Mian and Qamar (2008)

- What are the benefits of being in a network?
- Data on networks:
  - Authors have data on directors of all Pakistani companies, public and private
  - Define two firms as connected if they share a common director
  - Define two firms are in the same network if they can be linked through connected firms
  - Since they have the entire universe of firms, they can construct networks for the whole economy
- They find that there is one very large "super network"
  - Contains 5% of firms
  - But 66% of bank credit!

Khwaja, Mian and Qamar (2008)

- What are the benefits of being in a network?
- Data on networks:
  - Authors have data on directors of all Pakistani companies, public and private
  - Define two firms as connected if they share a common director
  - Define two firms are in the same network if they can be linked through connected firms
  - Since they have the entire universe of firms, they can construct networks for the whole economy
- They find that there is one very large "super network"
  - Contains 5% of firms
  - But 66% of bank credit!
- Empirical question:
  - What is the value of being in the super-network?

Constructing a network

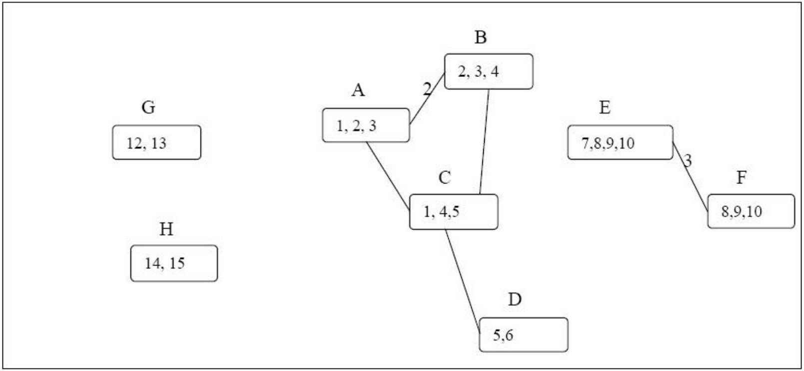
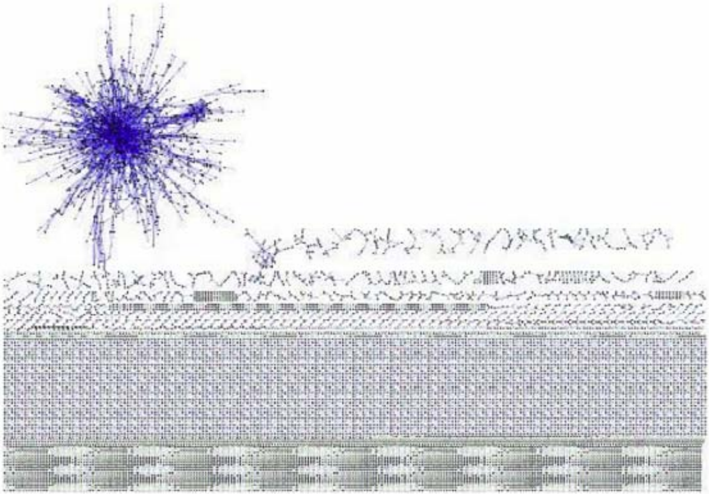


Figure I. Constructing Networks. This figure illustrates the hypothetical construction of a network. There are 8 firms in the example (A through H), and a total of 15 directors sitting on the board of these firms (labeled 1 through 15). Interlocked board linkages produce two distinct networks and two firms (G and H) that are not connected to anyone else. The largest network consists of firms A through D, where firms A, B and C are linked to each other directly and firm D is linked to firms A and B indirectly through its direct link with C. Thus firms in the same network may be linked to each other through chains of indirect links.

Networks in the economy



Appendix Figure I. Network structures in the entire economy. Firms are linked if they share a common director. The spatial positioning of various networks is in order of network size starting with the super-network.

Empirical strategy

- Compare super-network vs. non super-network firms.

Empirical strategy

- Compare super-network vs. non super-network firms.
  - Problem? Sign of bias?

Empirical strategy

- Compare super-network vs. non super-network firms.
  - Problem? Sign of bias?
- Do the same, but with firm fixed effects

Empirical strategy

- Compare super-network vs. non super-network firms.
  - Problem? Sign of bias?
- Do the same, but with firm fixed effects
  - Where does variation come from?
  - Problem? Sign of bias?

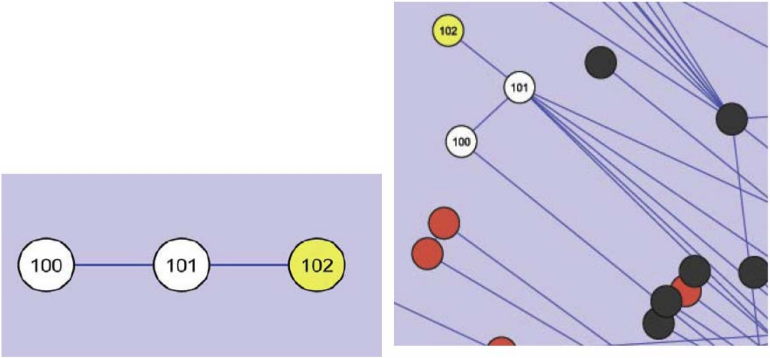
Empirical strategy

- Compare super-network vs. non super-network firms.
  - Problem? Sign of bias?
- Do the same, but with firm fixed effects
  - Where does variation come from?
  - Problem? Sign of bias?
- Empirical idea: use incidental firm entry and exits from the super-network

Empirical strategy

- Compare super-network vs. non super-network firms.
  - Problem? Sign of bias?
- Do the same, but with firm fixed effects
  - Where does variation come from?
  - Problem? Sign of bias?
- Empirical idea: use incidental firm entry and exits from the super-network
  - I.e., not whether your firm entered or exited the super-network, but whether another firm in your network entered or exited the supernetwork
  - Problem? Sign of bias?

Empirical strategy



Courtesy of Asim Ijaz Khwaja, Atif Mian, and Abid Qamar. Used with permission.

Results on Borrowing

- Estimate

$$Y_{it} = \alpha_i + \alpha_{kt} + \alpha_t + \gamma \Delta Y_{i,t-1} + \beta_1 ENTRY_{it} + \beta_1 ENTRY_{it} * Direct_i + \varepsilon_i$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
InNetwork	0.166***	0.184***	0.154***	0.177***	0.183***	0.128**	0.127**
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.059)	(0.059)
Lagged Loan Growth					0.012***		0.012***
					(0.002)		(0.002)
InNetwork * (Direct Entrant/Exitor)						0.126	0.126
						(0.085)	(0.085)
Fixed Effects	Basic	Expanded	Basic	Expanded	Expanded	Expanded	Expanded
Observations	286,034	286,034	12,053	12,053	286,034	286,034	286,034
R-squared	0.59	0.60	0.44	0.36	0.60	0.60	0.60

Results on Probability of Default

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
InNetwork	-1.728***	-1.632	-1.689***	-1.62***	-1.951***	-1.502***	-1.848***
	(0.350)	[0.351]***	(0.349)	(0.348)	(0.407)	(0.464)	(0.559)
Lagged Default Rate Growth					0.167***		0.167***
					(0.004)		(0.004)
InNetwork * (Direct Entrant/Exitor)						-0.284	-0.218
						-0.702	-0.807
Fixed Effects	Basic	Expanded	Basic	Expanded	Expanded	Expanded	Expanded
Observations	397,416	397,416	15,043	15,043	254,576	397,416	254,576
R-squared	0.86	0.86	0.1	0.002	0.86	0.86	0.86

Note: coefficients multiplied by 100

Mechanisms

	(1) Average Loan Size	(2) Total number of creditors	(3) %age credit from government banks	(4) %age credit from private banks	(5) New Credit Share from Neighbors' Lenders
InNetwork	0.139***	0.137***	-0.014***	0.025***	0.120***
	(0.041)	(0.018)	(0.003)	(0.005)	(0.005)
Constant	6.412***	1.041***	0.294***	0.51***	0.211***
	(0.021)	(0.004)	(0.002)	(0.002)	(0.005)
Fixed Effects	Expanded	Expanded	Expanded	Expanded	Expanded
Observations	286,034	286,034	286,034	286,034	30,065
R-squared	0.57	0.86	0.9	0.86	0.88

Bertrand, Mehta, and Mullainathan (2002)

- What is the downside of being in a network?

Bertrand, Mehta, and Mullainathan (2002)

- What is the downside of being in a network?
- Control rights over a firm and cash flow rights over firm's profits are not identical:
    - Control rights are awarded to whoever has a majority
    - Cash flow rights are awarded in proportion to ownership

Bertrand, Mehta, and Mullainathan (2002)

- What is the downside of being in a network?
- Control rights over a firm and cash flow rights over firm's profits are not identical:
  - Control rights are awarded to whoever has a majority
  - Cash flow rights are awarded in proportion to ownership
- With pyramid ownership structures, these can be totally separated:
  - Principal owns 51% of company i
  - Company i owns 51% of company $i + 1$
  - As $i \to \infty$ principal retains complete control but has 0 cash flow rights

Tunneling

- With these types of corporate structures, those with control rights have incentives to expropriate minority shareholders

Tunneling

- With these types of corporate structures, those with control rights have incentives to expropriate minority shareholders
- How?
  - Give loans to other firms in groups at artificially high/low interest rates
  - Sell assets to other firms in the group at artificially high/low prices
  - Etc

Tunneling

- With these types of corporate structures, those with control rights have incentives to expropriate minority shareholders
- How?
  - Give loans to other firms in groups at artificially high/low interest rates
  - Sell assets to other firms in the group at artificially high/low prices
  - Etc
- Why do we care?
  - If minority shareholders will be expropriated, means that business groups will have trouble attracting equity finance for their entities
  - This offsets the potential benefits of business groups discussed above

Tunneling

- With these types of corporate structures, those with control rights have incentives to expropriate minority shareholders
- How?
  - Give loans to other firms in groups at artificially high/low interest rates
  - Sell assets to other firms in the group at artificially high/low prices
  - Etc
- Why do we care?
  - If minority shareholders will be expropriated, means that business groups will have trouble attracting equity finance for their entities
  - This offsets the potential benefits of business groups discussed above
- Point of this paper is to detect tunneling

Methodology

- Idea: Consider external shock to predicted profits, and examine how actual profits respond to predicted profits
- Predictions:
  - Actual profits should respond less to predicted profits if firm is in a group
  - Response is smaller the lower the cash flow rights of the controlling firm
  - Controlling firm's profits should be more responsive to the bottom firm's shocks than their cash flow rights would imply
  - Response is greater if they have low cash flow rights
    - (this I don't see–seems to ignore actual profits)
  - Asymmetry: bottom firms profits are not sensitive to top firm's shocks
    - This distinguishes tunneling from insurance

Data

- Outcome: Profits Before Interest Depreciation and Taxes (PBIDT)
- Shocks: Average asset-weighted industry returns (excluding your firm)
  - Why exclude your firm?
- Cash flow rights:
  - Measure direct cash flow rights with several proxy variables:
    - Cash flow rights of directors (likely to be from the controlling group)
    - Cash flow rights of "other shareholders" (not directors, financial instituions, government bodies, corporate bodies, nor top fifty shareholders)
  - No measure of indirect cash flow rights (i.e., cash flow through intermediate firms)
    - Does this matter?

Regressions and Results

- Question 1: sensitivity to own shocks

$$\pi_{kt} = a + b\left(pred_{kt}\right) + c\left(cash_k * pred_{kt}\right) + dX_{kt} + \alpha_k + \alpha_t + \varepsilon_{kt}$$

TABLE II
SENSITIVITY TO OWN SHOCK: GROUP VERSUS STAND-ALONE
DEPENDENT VARIABLE: PROFIT BEFORE DIT

	(1)	(2)	(3)	(4)
Own shock	1.05	.10	−4.58	−5.10
	(.02)	(.05)	(.48)	(.47)
Own shock* group	**−.30**	**−.30**	**−.26**	**−.27**
	(.02)	(.02)	(.02)	(.02)
Ln assets	.16	2.98	−.33	2.47
	(.32)	(.34)	(.33)	(.34)
Own shock* ln assets	—	.10	—	1.0
		(.00)		(.01)
Own shock* year of incorp.	—	—	.003	.003
			(.000)	(.000)
Sample size	18600	18600	18588	18588
Adjusted R^2	.93	.93	.93	.93

a. *Data Source:* Prowess, Centre for Monitoring Indian Economy, for years 1989–1999. All monetary variables are expressed in 1995 Rs. crore, where crore represents 10 million. Sample includes both stand-alone and group firms.
b. All regressions also include year fixed effect and firm fixed effects.
c. Standard errors are in parentheses.

Regressions and Results

- Question 2: sensitivity to amount of director equity

TABLE III
SENSITIVITY TO OWN SHOCK BY DIRECTOR AND OTHER OWNERSHIP
DEPENDENT VARIABLE: PROFIT BEFORE DIT

Panel A: Director equity

	Sample: Groups (1)	Groups (2)	Stand-alones (3)	Stand-alones (4)
Own shock	.713	−5.075	1.058	−4.316
	(.009)	(.742)	(.006)	(.518)
Own shock * director equity	**.025**	**.030**	**.004**	**.019**
	(.003)	(.003)	(.001)	(.001)
Ln assets	.052	4.261	−.590	1.568
	(.733)	(.807)	(.176)	(.178)
Own shock * ln assets	—	.118	—	.201
		(.008)		(.006)
Own shock * year of incorp.	—	.002	—	.002
		(.000)		(.000)
Sample size	7521	7510	11079	11078
Adjusted R^2	.92	.93	.95	.96

Regressions and Results

- Question 5: Is there asymetry, i.e., do profits move towards the 'top' firm in the chain?

TABLE V

SENSITIVITY TO GROUP SHOCK BY LEVEL OF DIRECTOR OWNERSHIP IN GROUP

DEPENDENT VARIABLE: PROFIT BEFORE DIT

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Level in group:	Lower ⅔	Top ⅓	Below topmost firm			Topmost firm		
Own shock	.62	.89	.63	.63	.63	1.01	1.01	1.01
	(.01)	(.02)	(.01)	(.01)	(.01)	(.02)	(.02)	(.02)
Group shock	.013	.010	.012	—	—	.020	—	—
	(.002)	(.002)	(.001)			(.008)		
Shock below 66th pctile (director equity)	—	—	—	.015	—	—	.032	—
				(.002)			(.012)	
Shock above 66th pctile (director equity)	—	—	—	.003	—	—	.007	—
				(.006)			(.018)	
Shock below 33rd pctile (other ownership)	—	—	—	—	−.000	—	—	−.013
					(.004)			(.025)
Shock above 33rd pctile (other ownership)	—	—	—	—	.017	—	—	.034
					(.002)			(.011)
Sample size	4905	2616	5780	5780	5780	1741	1741	1741
Adjusted R^2	.90	.95	.90	.97	.97	.97	.97	.97

a. *Data Source:* Prowess, Centre for Monitoring Indian Economy, for years 1989–1999. All monetary variables are expressed in 1995 Rs. crore, where crore represents 10 million.

b. Firms are separated into different "Level in group" based on their within-group level of director equity. For example, "Topmost Firm" are the set of firms that have the highest level of director ownership in their group.

c. Also included in each regression are the logarithm of total assets, year fixed effects, and firm fixed effects.

d. Standard errors are in parentheses.

Family firms

- Many firms are by family members of the original founder.
- A priori, this seems inefficient: why would we think that managerial talent is hereditary? Shouldn't the market find a better manager?
- Why might this be?

Burkart, Panunzi, and Shleifer (2003)

- Tunneling!

Burkart, Panunzi, and Shleifer (2003)

- Tunneling!
- Assume no superior manager has resources to buy firm outright

Burkart, Panunzi, and Shleifer (2003)

- Tunneling!
- Assume no superior manager has resources to buy firm outright
- Then:
  - If shareholder protections are strong, then you can sell all your stock in the company, and it is run with diversified ownership.
  - If shareholder protections are intermediate, you sell some stock but continue to be a large shareholder, and monitor the professional manager to limit expropriation.
  - If shareholder protections are very weak, so even a manager can expropriate a large shareholder, you retain control within the family.

Perez-Gonzalez (2006)

- What is the impact of inherited management on firm performance? Is it actually negative (as above model suggests)?

Perez-Gonzalez (2006)

- What is the impact of inherited management on firm performance? Is it actually negative (as above model suggests)?
- Idea:
  - Look at firms that were initially controlled by a family, and where there was a CEO succession
  - Compare stock returns for those firms that announce family member will be new CEO with those that announce external new CEO
  - Similarly compare change in actual profits before and after new CEO takes over

Stock-market event studies

- Stock market event studies:
  - Basic idea: efficient markets hypothesis implies that the full long run value of new information on a firm is incorporated in the stock price immediately
  - So the change in a stock's price right around the time of new information tells you the value of that new information

Stock-market event studies

- Stock market event studies:
  - Basic idea: efficient markets hypothesis implies that the full long run value of new information on a firm is incorporated in the stock price immediately
  - So the change in a stock's price right around the time of new information tells you the value of that new information
- Development examples:
  - Fisman (2001) studies effect of Suharto's health on connected firms to determine the value of political connections
  - Guidolin and La Ferrara (2007) studies impact of conflict shocks on mineral firms to detect illegal diamond trade

Stock-market event studies

- Estimation:
  - Estimate a market model to find "abnormal returns" for a firm, i.e. take the residuals from

$$r_f = \alpha + \beta r_m + \varepsilon_f$$

  - Define a window around the event e.
  - Then estimate average abnormal returns during the event window e and test the null that they are equal to 0.
- What do we learn from these models? When might they be reasonable? When might they not be reasonable?

Results

TABLE 3—CUMULATIVE ABNORMAL RETURNS AROUND SUCCESSION ANNOUNCEMENTS

Firms and event-window	All	Type of succession: Family	Type of succession: Unrelated	Difference	Mann-Whitney $\|z\|$-values
	(1)	(2)	(3)	(4)	(5)
All CEO successions (t_0, t_{+2})	0.0100	−0.0018	0.0167	−0.0184	1.265
	(0.0044)	(0.0071)	(0.0055)	(0.0089)	
	[335]	[122]	[213]		
All CEO successions (t_0, t_{+5})	0.0096	−0.0016	0.0160	−0.0176	1.585
	(0.0047)	(0.0068)	(0.0063)	(0.0093)	
	[335]	[122]	[213]		
Successions reported as "retirements" (t_0, t_{+2})	0.0096	−0.0020	0.0165	−0.0185	1.121
	(0.0049)	(0.0083)	(0.0060)	(0.0103)	
	[260]	[97]	[163]		

Results

- Also examines changes in accounting profits

Years	All	Family	Unrelated	Difference
	(1)	(2)	(3)	(4)
Number of CEO transitions	335	122	213	
A. Operating return on assets (OROA)				
(*3-year average after*) − (*3 year average before*)	−0.0055	−0.0188	0.0021	−0.0209
	(0.0039)	(0.0059)	(0.0050)	(0.0077)
B. Industry adjusted OROA				
(*3-year average after*) − (*3 year average before*)	0.0022	−0.0114	0.0100	−0.0213
	(0.0040)	(0.0063)	(0.0051)	(0.0081)
C. Industry and performance adjusted OROA				
(*3-year average after*) − (*3 year average before*)	0.0071	−0.0059	0.0146	−0.0205
	(0.0037)	(0.0056)	(0.0049)	(0.0074)
$(t = -1) - (t = -3)$	−0.0121	−0.0169	−0.0093	−0.0076
	(0.0041)	(0.0080)	(0.0046)	(0.0093)
$(t = +3) - (t = -1)$	0.0120	−0.0003	0.0191	−0.0194
	(0.0052)	(0.0097)	(0.0059)	(0.0113)

Concluding thoughts

- Firms are important engines of economic growth
- Problems with contracting and credit lead to unusual corporate stuctures, with some benefits but also some costs
- But I think there's much more about firms that hasn't been explored much.
- Some things I think are interesting:
    - Business clusters
    - Branding
    - Endogenous adoption of technology
    - Internal firm capital markets
    - Political capture
    - Firm behavior

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

14.771: Technology Lecture 1

Ben Olken

October 2008

Outline

- Technology Adoption (through the lens of agriculture)
  - Does it matter? Why doesn't everyone adopt? What's the big problem here?
  - Learning
  - Savings and other problems
- Other issues in technology
  - How learning from neighbors can get you the wrong answer
  - How technology can affect markets
  - Appropriate technology

What's the problem? Why doesn't everyone adopt?

- Duflo, Kremer and Robinson (2008)
- Setting:
  - Maize farming in Kenya
  - Technology is fertilizer ("top-dressing", which is fertilizer applied after plant has germinated and probability it will grow to fruition is high)
  - Farmer adds fertilizer 2 months after planting
  - "Return" is realized 7 months later, when farmer can consume extra maize the produced rather than buy it at the market price
- Design:
  - Randomized experiment where farmers are randomized into different levels of fertilizer use or control

Results

- Key result: potential for very high returns, but only if you get the amounts right

Table 1: Returns to Fertilizer

	Mean (1)	Median (2)	Std. Error (3)	Obs. (4)
Panel A. 1/4 Teaspoon Top Dressing Fertilizer				
Percentage Increase in Yield	28.1	8.9	6.8	112
Rate of Return Over the Season	4.8	-27.7	38.8	112
Annualized Rate of Return (at the Mean and Median)	8.4	-42.6		112
Panel B. 1/2 Teaspoon Top Dressing Fertilizer				
Percentage Increase in Yield	47.6	24.3	6.1	200
Rate of Return Over the Season	36.0	23.9	16.9	202
Annualized Rate of Return (at the Mean and Median)	69.5	44.4		202
Panel C. 1 Teaspoon Top Dressing Fertilizer				
Percentage Increase in Yield	63.1	30.6	8.2	273
Rate of Return Over the Season	-10.8	-16.9	8.4	274
Annualized Rate of Return (at the Mean and Median)	-17.8	-27.3		274
Panel D. Full Package Recommended by Ministry of Agriculture				
Percentage Increase in Yield	90.6	48.7	15.4	82
Rate of Return Over the Season	-38.9	-49.4	10.4	85
Annualized Rate of Return (at the Mean and Median)	-48.2	-59.7		85

Note: A/B/C vs. control were done in different seasons

Technology adoption

- Heterogeneity in returns suggests learning is important, not just about whether to adopt, but how to adopt
- Several ways to think about technology adoption:
- Learning
  - About (constant) returns to technology
  - About (constant) appropriate use of technology
  - About (idiosyncratic) appropriate use of technology
- Deciding whether to adopt
  - Risk aversion
  - Credit constraints
  - Time-consistent preferences

Learning: the empirical identification challenge

- General problem: in cross section, correlated shocks cannot be distinguished from learning
- Consider the simplest case:

$$y_{ig} = \alpha + \beta \overline{y}_g + u_g + \varepsilon_i$$

  where u is an unobserved common shock. Since u is unobserved, we cannot identify β.
- Some options for identification we will discuss:
  - Use panel data and exploit time structure (Foster and Rosenzweig)
  - Distinguish between learning group and information group so $\overline{y}_g$ and u_s refer to different groups (Conley and Udry)
  - Use experiments to shock one individual only so we know there are no common shocks (Duflo, Kremer, and Robinson)

Foster and Rosenzweig (1995)

- Overview:
- Model of agricultural technology adoption
- In the model, optimal use of technology is idiosyncratic, so farmers need to experiment to learn how to use the technology in a way appropriate for local conditions
- Farmers can learn both from their own experience and from the experience of their neighbors
- This predicts that farmers under-experiment relative to the social optimum, so adoption is too slow
- Test model using panel data from India

Model setup

- Optimal use of technology is

$$\widetilde{\theta}_{ijt} = \theta^* + u_{ijt}$$

  where θ^* is mean (unknown) optimal use and u_{ijt} is iid error term with variance σ_u^2.
- Farmers have priors over θ^* that are $N\left(\theta_{j0}, \sigma_{\theta j0}^2\right)$.
- HYV yield for parcel i:

$$\eta_a + \eta_h - \eta_{ha}\frac{i}{A_j} - \left(\theta_{ijt} - \widetilde{\theta}_{ijt}\right)^2$$

  where a is alternative, h hybrid, A_j is total number of plots, i is plot number
- Expected profits for a farmer are therefore

$$E\pi_{jt} = \left(\eta_h - \eta_{ha}\frac{H_j}{2A_j} - \sigma_{\theta jt}^2 - \sigma_u^2\right) H_j + \eta_a A_j + \mu_j + \varepsilon_{pjt}$$

Learning

- Farmers update their priors about the optimal way to use the technology (θ) using Bayes' rule
- This implies that

$$\sigma^2_{\theta jt} = \frac{1}{\rho + \rho_0 S_{jt} + \rho_v \overline{S}_{-jt}}$$

where ρ's are "precisions" of information (inverse of variances):

$$\begin{aligned} \rho &= \frac{1}{\sigma^2_{\theta 0}} \\ \rho_0 &= \frac{1}{\sigma^2_u} \\ \rho_v &= \frac{n}{(\sigma^2_u + \sigma^2_k)} \end{aligned}$$

and where S_{jt} is own experience, n is number of neighbors, and $\overline{S}_{-jt}$ is average neighbor experience (you observe neighbor with more noise)

Technology Adoption

- Farmer's choice variable is H_{jt}, i.e., how many plots to plant HYVs in this year
- Farmer chooses H_{jt} to maximize:

$$V_{jt} = \max_{H_{jx}} E_t \sum_{x=t}^{T} \delta^{x-t} \pi_{jx}$$

or equivalently in Bellman form:

$$\begin{aligned} V_{jt} &= \max_{H_j} \left(\eta_h - \eta_{ha} \frac{H_j}{2A_j} - \frac{1}{\rho + \rho_0 S_{jt} + \rho_v \overline{S}_{-jt}} - \sigma_u^2 \right) H_j \\ &\quad + \eta_a A_j + \mu_j + \delta V_{jt+1} \end{aligned}$$

- Planting decisions therefore depend (positively) on past history of own and neighbors' planting, (positively) on expectations about own future planting, and (ambiguous, maybe negative) on expectations about neighbors' planting

Technology Adoption

- Taking derivatives with respect to H_j the first order condition is that:

$$\eta_h - \eta_{ha}\frac{H_j}{A_j} - \frac{1}{\rho + \rho_0 S_{jt} + \rho_v \overline{S}_{-jt}} - \sigma_u^2 = -\delta \frac{\partial V_{t+1}}{\partial S_{jt}}$$

- Note that at optimum interior solution, current period marginal benefit of H is negative – i.e., you do some extra experimentation to gain knowledge usable in the future

Restrictions/Implications

- Ratio of marginal impact of own and neighbors experience is a time-invariant constant since relative precision of additional information gained by each does not change over time – i.e,

$$\frac{\frac{\partial \pi_{jt}}{\partial S_{jt}}}{\frac{\partial \pi_{jt}}{\partial \overline{S}_{-jt}}} = \frac{\rho_0}{\rho_v}$$

- Value of additional own and neighbor's information diminish over time as long as HYV use is positive, and at the same rate, i.e.,

$$\frac{\frac{\partial(\pi_{jt+1}/H_{jt+1})}{\partial S_{jt+1}}}{\frac{\partial(\pi_{jt}/H_{jt})}{\partial S_{jt}}} = \frac{\frac{\partial(\pi_{jt+1}/H_{jt+1})}{\partial \overline{S}_{-jt+1}}}{\frac{\partial(\pi_{jt}/H_{jt})}{\partial \overline{S}_{-jt}}} = \frac{\left(\rho + \rho_0 S_{jt} + \rho_v \overline{S}_{-jt}\right)^2}{\left(\rho + \rho_0 S_{jt+1} + \rho_v \overline{S}_{-jt+1}\right)^2} < 1$$

- If no learning from neighbors $(\rho_v = 0)$, then neighbor's assets (A_{-j}) do not affect farmer's decisions
- Effect of neighbor's assets that predict HYV planting could be negative, although own effects of assets are positive

Estimation

- Recall profit function (augmented to include education, denoted E)

$$\pi_{jt} = \left(\eta_h - \eta_{ha}\frac{H_j}{2A_j} - \frac{1}{\rho + \rho_0 S_{jt} + \rho_v \overline{S}_{-jt}} - \sigma_u^2 + \eta_{he}E_j\right) H_j + \eta_a A_j + \mu_j + \varepsilon_{pjt}$$

- Taking linear approximation yields

$$\pi_{jt} = \left(\eta_h' + \beta_{0t}S_{jt} + \beta_{vt}\overline{S}_{-jt} + \eta_{he}E_j\right) H_j + \eta_a' A_j + \mu_j + \varepsilon_{pjt}$$

where

$$\beta_{ot} = \frac{\rho_o}{\rho + (\rho_o + \rho_v) S_t}, \beta_{vt} = \frac{\rho_v}{\rho + (\rho_o + \rho_v) S_t}$$

where S_t is some average level of S around which we take approximations.

- How does this differ from what you might have written down from reduced form perspective?
    - Main effects of S_{jt} and S_{-jt}? Which makes more sense?

Estimation notes

- Take first differences to remove fixed effect, and retain both current and lagged S variables. This yields:

$$\begin{aligned}\Delta\pi_{jt} &= \eta'_h \Delta H_j + \beta_{0t+1} S_{jt+1} H_{jt+1} + \beta_{vt+1} \overline{S}_{-jt+1} H_{jt+1} \\ &\quad - \beta_{0t} S_{jt} H_{jt} - \beta_v \overline{S}_{-jt} H_{jt} - \eta_{he} E_j \Delta H_{jt} + \eta'_a \Delta A_{jt} + \Delta\varepsilon_{pjt}\end{aligned}$$

- By allowing coefficients β_0 and β_v to differ over time, they can test whether learning changes over time
- Estimate using instrumental variables.
  - What are the concerns?
    - Weather, pests: if some component of profitability is known ex-ante and affects HYV adoption decision, you could get bias.
    - Lagged profit shocks can affect HYV adoption through learning
  - Instruments are inheritance of assets and lags of ΔA and H (i.e., use levels of assets and lags to instrument for changes)
    - Are these good instruments?
- Adoption (ΔH_{jt}) regressions are similar

Results-Specification Check

- Check: village experience with HYV should not affect profitability of non-HYV farmers. Not true in cross-section but true in panel.

Cross-Sectional and Panel Estimates of Profit Function for Farmers not Using HYVs

	OLS (N = 1,536)	*Fixed effects (N = 1,277)*	*Instrumental variables fixed effects (N = 1,277)*	
	(1)	*(2)*	*(3)*	*(4)*
Village experience	.137 (1.84)	-.187 (.654)	-.246 (.804)	-.240 (.784)
Initial period village experience				.166 (.514)
Equipment	.085 (1.29)	.597 (2.11)	2.94 (2.90)	2.90 (2.85)
Irrigation assets	.162 (7.68)	.050 (.691)	.425 (2.00)	.440 (2.06)
Animals	.657 (17.9)	-.377 (2.30)	-1.74 (4.16)	-1.76 (4.20)
Primary schooling (x 10^2)	1.77 (2.01)	--	--	--
Irrigated land	.018 (7.01)	--	--	--
Unirrigated land	.032 (9.34)	--	--	--
House	.026 (3.41)	--	--	--

Note: All variables are treated as endogenous for instrumental variables, fixed-effect estimates. Instruments include inherited assets, lagged asset flows, lagged profits, lagged village HYV use, and weighted averages of these variables by village. Absolute asymptotic t-ratios derived from Huber standard errors are in parentheses.

Figure by MIT OpenCourseWare.

Results-Profits

- Own and neighbor experience matters (but own matters much more)_, experience declines over time
- Ratio of decline over time similar for own and neighbor information

Determinants of Farm Profits from HYV Use (N = 450)

HYV effects	*Linear approximation*			*Structural estimates: Nonlinear instrumental variables fixed effects*
	Instrumental variables fixed effects		*Constrained instrumental variables fixed effects*	
	(1)	*(2)*	*(3)*	*(4)*
β_α (x 10^5)	.170 (2.13)	.293 (2.54)	.187 (1.88)	--
$\beta_{\alpha-1}$ (x 10^5)	.754 (2.47)	1.05 (2.18)	--	--
β_α (x 10^5)	--	.349 (2.16)	.341 (2.63)	--
$\beta_{\alpha-1}$ (x 10^5)	--	1.93 (2.64)	--	--
λ_μ	--	--	4.33 (10.6)	--

Figure by MIT OpenCourseWare.

Results-Adoption

- Neighbors assets negatively affect your adoption – which they interpret as evidence of free riding on experimentation

Farm equipment: neighbor (x 10^{-4})	--	-.0878 (.34)	-.0194 (.06)
Farm animals: neighbor (x 10^{-4})	--	-.995 (2.08)	-.948 (1.85)
Irrigation assets: neighbor (x 10^{-4})	--	-2.12 (3.58)	-2.07 (3.38)
Trend (x 10^{-2})	3.85 (2.54)	4.04 (2.65)	4.07 (2.53)

Figure by MIT OpenCourseWare.

Discussion

- In general, it is very hard to separate peer effects (learning) from common shocks
- In this paper, they do this in two ways:
  - Looking at lag values
  - Instrumenting using levels (in a difference equation) and lags
  - Are these convincing?
- Does learning from others speed up or slow down technology adoption?
  - Two effects: benefits from learning, but free riding on neighbors
  - On net simulations imply profits improve with spillovers, but time to adoption is slower

Conley and Udry (2005)

- Setting:
  - Pineapple growers in Ghana
  - Learning whether and how much fertilizer to use from the experience of neighbors
- Similar idea to F&R: technology is local, so you need to see how it works locally to use it correctly

Conley and Udry vs. Foster and Rosenzweig

- Data improvements
  - Data on who you discuss farming with, so we can construct the people who you learn from better than village level aggregates in F&R
  - Moreover, network information allows them to distinguish between spacial correlation (common shocks) and who you learn from
- Sharper empirics
  - In F&R, the idea is you learn about optimal input level from neighbor's experience but this is unobserved
  - In C&Y, they observe input choices of you and neighbors. They then test whether you change your inputs in response to new information about productivity of neighbors.
  - This is the same prediction as the F&R model; C&Y they just have the data to look at it directly

Empirical setup

- Two categories: use fertilizer ($x > 0$) or not ($x = 0$)
- For each level of x, calculate expectations of profits based on past realizations of profits for neighbors. Denote this $\widehat{E}_{it}\left[\pi_{k,t+1}\left(x_{kt}, w_{kt}\right)\right]$ where w measures growing conditions and k is a similar set of plots.
- "Good news" if

$$d_{i,k,t} = \pi_{k,t+1}\left(x_{kt}, w_{kt}\right) > \widehat{E}_{it}\left[\pi_{k,t+1}\left(x_{kt}, w_{kt}\right)\right]$$

- Index of good news defined as number of neighbor's plants with same input type for which good news was reported, expressed as share of total neighbor's plants ever seen

$$G_{i,t_1}\left(x = x_{i,t_0}\right) = \frac{1}{TotalPlants_{i,t_t}} \sum_{k \in N_i} 1\left\{x_{i,t_0} = x_{k,t_0}\right\} d_{i,k,t_0} Plants_{k,t_0}$$

- Define good news for alternative input choices and bad news for same and alternate input choice analogously

Empirical setup

- Similarly, define
  - M as where (relative to x_0) the good news is. You should therefore update in the direction of M.
  - Γ as difference between what you and your geographic neighbors do
    - Not quite analogous – we would probably have liked it to be exactly the same to really tease out geography vs. information sets, but so be it
- Probability of change regression

$$\Pr\{\Delta x_{i,t} \neq 0\} = \Lambda \left[\begin{array}{c} \alpha_1 G_{it}(x = x_{i,t-1}) + \alpha_2 G_{it}(x \neq x_{i,t-1}) \\ +\alpha_4 B_{it}(x = x_{i,t-1}) + \alpha_4 B_{it}(x \neq x_{i,t-1}) \\ +\alpha_5 Geog + z'_{I,T}\alpha_6 \end{array} \right]$$

- Directional change regression

$$\Delta x_{i,t} = \beta_1 M_{i,t} + \beta_2 \Gamma_{i,t} + z'_{i,t}\beta_3 + \varepsilon$$

Results

Table 4: Predicting the Change in Input Use

Dependent Variable: Indicator of a Change in Per Plant Fertilizer Use

	A	B
Good News at Lagged Fertilizer Use	-0.13 [1.19]	0.05 [1.16]
Good News at Alternative Fertilizer Use	0.18 [0.97]	0.37 [1.02]
Bad News at Lagged Fertilizer Use	12.32 [3.72]	14.41 [4.63]
Bad News at Alternative Fertilizer Use	-2.98 [1.91]	-4.22 [2.07]
Average Absolute Deviation from Geographic Neighbors' Fertilizer Use	0.49 [0.13]	0.49 [0.14]

Results

Table 5: Predicting Innovations in Input Use, Differential Effects by Source of Information

Dependent Variable: Innovation in Per Plant Fertilizer Use

	A	B	C	D	E	F
Index of Inputs on Successful Experiments (M)	0.99					
	[.16]					
M * Inexperienced Farmer		1.09				
		[0.22]				
M * Experienced Farmer		0.10				
		[0.32]				
Inexperienced Farmer		4.01	4.20	4.22	4.19	4.12
		[2.62]	[2.66]	[2.65]	[2.65]	[2.77]
Index of Experiments by Inexperienced Farmers			-0.13			
			[0.37]			
Index of Experiments by Experienced Farmers			1.02			
			[0.17]			
Index of Exper. by Farmers with Same Wealth				1.03		
				[0.18]		
Index of Exper. by Farmers with Different Wealth				-0.41		
				[0.32]		
Index of Experiments on Big Farms					1.10	
					[0.14]	
Index of Experiments on Small Farms					0.89	
					[0.18]	
Index of Exper. by Farmers with Same Soil						1.04
						[0.16]
Index of Exper. by Farmers with Different Soil						0.91
						[0.19]

Discussion

- These results provide more direct confirmation for the Foster & Rosenzweig story:
  - People adjust their fertilizer use based on neighbor's experience, particularly bad news from n
  - Effects particularly strong for inexperienced farmers.
  - People also pay attention more to farmers with similar wealth and more experience

Duflo, Kremer and Robinson (2006)

- Same experiments on fertilizer use in Kenya discussed at beginning of lecture
- Several treatments to investigate learning:
  - Shock farmers with information and test inputs (fertilizer) in period 1 and look at impact over time ("demonstration plot")
  - Shock farmers with test inputs only (fertilizer) in period 1 and look at impact over time ("starter kit")
  - Ask demonstration plot farmers (in treatment and control) to name people with whom they discuss agriculture. Then randomly invite one of those people to attend discussion.
- Key dependent variable: did you adopting fertilizer in subsequent years

Results

- Farmers in demonstration plot continue to use over time, but effect is only 10-17 percentage points. Suggests other factors may matter

Table 2: Adoption for Farmers Participating in Demonstration Plot

	1 season later		*2 seasons later*		*3 seasons later*	
Panel A. All Farmers	(1)	(2)	(3)	(4)	(5)	(6)
	fertilizer	*any treatment*	*fertilizer*	*any treatment*	*fertilizer*	*any trea*
Demonstration Plot Farmer	0.107	0.107	0.089	0.061	0.101	0.0
	(0.039)***	(0.042)**	(0.044)**	(0.046)	(0.050)**	(0.05
Observations	580	577	523	521	450	447
Panel B. Only Farmers with at least 3 seasons of Adoption Data						
	(1)	(2)	(3)	(4)	(5)	(6)
	fertilizer	*any treatment*	*fertilizer*	*any treatment*	*fertilizer*	*any trea*
Demonstration Plot Farmer	0.169	0.172	0.113	0.065	0.112	0.07
	(0.050)***	(0.056)***	(0.054)**	(0.059)	(0.054)**	(0.06
Observations	371	364	371	364	371	364

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Regressions control for school.

Courtesy of Esther Duflo, Michael Kremer, and Jonathan Robinson. Used with permission.

Results

- Starter kits half as effective and effect diminishes over time – suggests external information is important

Table 3: Adoption for Farmers Offered Starter Kits

	1 season later		2 seasons later	
	(1)	(2)	(3)	(4)
	fertilizer	*any treatment*	*fertilizer*	*any treatmer*
Starter Kit Farmer	0.051	0.063	0.019	0.021
	(0.029)*	(0.029)**	(0.029)	(0.029)
Observations	1045	1042	1060	1059

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Regressions control for school.

Results

- Those invited to discuss results have similar effects as those who had experiment on own plot, although diminishes more rapidly
  - Maybe because without agricultural extension, they can't adapt fertilizer use properly to their own plots? We don't know.
- No spillovers on other neighbors.
  - Some evidence people in this area don't discuss farming much - only 46% of farmers can state what neighbors planted

Table 4: Adoption for Agricultural Contacts

	1 season later		*2 seasons later*		*3 seasons later*	
Panel A. OLS	(1)	(2)	(3)	(4)	(5)	(6)
	fertilizer	*any treatment*	*fertilizer*	*any treatment*	*fertilizer*	*any treatment*
Invited Agricultural Contact	0.093	0.103	0.031	0.054	-0.014	-0.022
	(0.055)*	(0.059)*	(0.061)	(0.065)	(0.058)	(0.064)
Uninvited Agricultural Contact	0.002	-0.042	-0.015	-0.004	-0.015	-0.013
	(0.035)	(0.038)	(0.038)	(0.041)	(0.037)	(0.040)
Observations	708	706	580	580	557	556
B. Panel B: 2SLS	(1)	(2)	(3)	(4)	(5)	(6)
	fertilizer	*any treatment*	*fertilizer*	*any treatment*	*fertilizer*	*any treatment*
Came to Treatment (instrumented	0.212	0.236	0.072	0.126	-0.03	-0.049
with Invited Agricultural Contact)	(0.127)*	(0.138)*	(0.142)	(0.151)	(0.129)	(0.141)
Uninvited Agricultural Contact	-0.002	-0.047	-0.017	-0.007	-0.014	-0.012
	(0.037)	(0.040)	(0.039)	(0.042)	(0.037)	(0.041)
Observations	708	706	580	580	557	556

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Concluding thoughts

- In general the results suggest that an important part of technology adoption is learning how to use it, not just whether to use it. There are therefore important returns to experience and spillovers
- Spillovers probably somewhat context dependent: depends how much you talk to neighbors
- Examples from agriculture, but this may be more general
  - Thought experiment: think about moving from using Fortran to Stata.
  - May be some negative returns up front since you don't know how to use Stata and you need to figure it out; your first attempts may not work
  - Easier to learn how to use it if you can ask your friends how to use it, get examples of their code, etc.
- Results from Kenya experiments suggest that there may be important other facts as well: more next time on this

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

14.771: Technology Lecture 2

Ben Olken

November 2008

Outline

- More on technology adoption:
  - Some pitfalls of learning: herd behavior
  - Savings and other constraints on technology adoption
- How technology can affect markets
- Other issues in technology
  - Appropriate technology
  - International technology transfer

Banerjee (1992)

- "A Simple Model of Herd Behavior" – we'll look at the very simple version
- Suppose there are two options, A and B. In the paper, they are restaurants, could also be ways of using a technology, investments, or whatever.
- One option is better than the other. If you choose the good restaurant you get return y; if bad restaurant you get return 0.
- Common priors over which is better. Suppose prior probability A is better is α.
- Each person receives a signal about which is better. Signal is correct with probability β.
- People move in sequence. You observe people's choices, but not their private signal.

Equilibrium in the Simple Model

- What happens?
- Person 1 follows signal.
- Person 2:
  - Observes their own signal, person 1's choice, and the common prior
  - Chooses whichever option has higher posterior probability
- Herd behavior:
  - Suppose person 1's signal matches the prior, and person 2 gets the opposite signal
  - Since both signals are of the same quality, person 2 has no information except the prior.
  - So person 2 ignores the private signal and chooses the option with higher prior.
  - By induction so will everyone else
- So everyone can end up on the wrong outcome!

Sources of inefficiency

- Why?
  - The key thing is that information can be lost – people's choices are not sufficient statistic for all the information that has been revealed.
- How would we prevent inefficient herding?
- How would model differ if:
  - Everyone moved at the same time
  - There were multiple discrete choices
  - There were a continuum of possible choices
    - Answer: depends on loss function. If quadratic loss like F&R, you don't get inefficiency, because your choice is a sufficient statistic for all previous information. If discrete gain from getting the right answer, you can continue to get inefficiency
- Key point: critical to learning is the precise nature of information revelation. Learning from others can be good, but the key is to ensure you don't get trapped in the wrong outcome.

Credit and hyberbolic discounting

- Duflo, Kremer and Robinson (2006), maize in Kenya
- Interviews with farmers suggested that one reason only 10%-17% of farmers in demonstration plots took up fertilizer themselves was "they didn't have the money"
- Could normal credit constraints be the problem?
- Given that farmers have cash right after the last harvest, seems like normal credit constraints may not be the problem – may be that they are not good at saving money for harvest

Present-biased preferences

- Traditional preferences

$$u\left(c_t\right) = \sum_{k=t}^{T} \delta^{k-t} v\left(c_k\right)$$

- Present-biased preferences (also called "hyberbolic discounting", see Loewenstein/Prelec, Laibson, Rabin and others) capture they idea that individuals may discount the entire future more than they discount any future period relative to the previous one

$$u\left(c_t\right) = v\left(c_t\right) + \beta \sum_{k=t+1}^{T} \delta^{k-t} v\left(c_k\right)$$

  $\beta < 1$ implies 'present-bias'
- Key implication: with $\beta < 1$, people prefer 100 dollars today to 110 in a month, but would prefer 110 in two months to 100 in one month. Such preferences would violate normal (exponential) discounting

Present-biased preferences

- These preferences are not time-consistent. Today I would like to start a saving plan tomorrow, but tomorrow I would like to put it off until a day later.
- Two ways of thinking about these preferences: naiive and sophisticated. Sophisticated people are aware of these preferences; naiive people are not.
- Sophisticated people have a demand for commitment: today I would like to place restraints on my future self.

Model of fertilizer adoption

- Consider a 3-period model
  - Period 1, inherit income from previous harvest
  - Period 2, plant. Chose to use fertilizer on share γ of land. Receive no income.
  - Period 3, receive income from next harvest

$$y_3 = \gamma y_H + (1 - \gamma) y_L$$

- Can purchase fertilizer in period 1 or 2, with small utility cost that is paid in the period when it is purchased. (hassle)
- Utility function in period 1

$$u(c_1) + \beta(u(c_2) + u(c_3)) + F * D_1 + \beta F * D_2$$

where D_j is 1 if fertilizer purchased in period j and F is utility cost

Model of fertilizer adoption

- What will a naiive farmer do?
- What will a sophisticated farmer do?

Empirical tests

- Randomize farmers into the following treatments:
  - Farmer is visited by agent at harvest and offered option to buy fertilizer then. Take it or leave it.
  - Farmer is visited by agent before harvest, and asked when person should return to sell him fertilizer. When returns it is as above.
  - Control
- Also examined subsidizing price of fertilizer by 50%

Results

Results

- Program is taken up by 30%-40% of farmers and increases adoption by 10%-12%
- Effect of visiting early is comparable to 50% reduction in price – quite substantial effect
- When farmers are given ex-ante choice of when to come back, many choose to have the person come back right after harvest: suggests some amount of sophistication is present. Total effect is similar to main program: suggests not just impulse buy
- Bottom line: these savings / procrastination stories may be important, and magnitude is as large as a 50% reduction in price

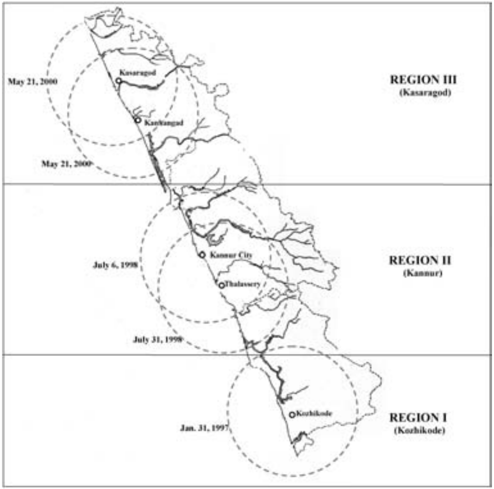
Other issues

- So far everything we've discussed has been about technology adoption.
- Switching gears.... Three more topics in brief.
  - High tech and development: does it matter?
  - Appropriate technology for developing countries
  - International technology transfer

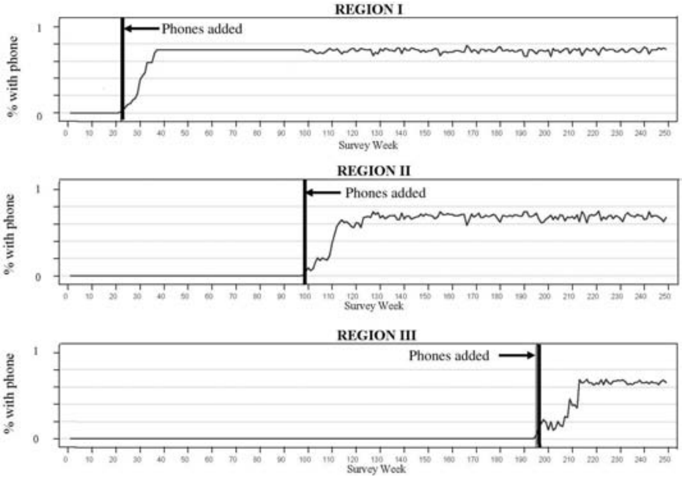
Jensen (2008)

- Setting:
    - Fisherman in Kerala, India
    - Fish markets are located every 15km or so up and down the coast of Kerala
    - Travel time means each fisherman only has time to bring catch to one market
    - No storage – fish caught that day must be sold that day
- Experiment:
    - Introduction of cell phones along the Kerala coast in three phases
    - Cell phones mean that fisherman can call/SMS ahead while still at sea to determine which market to go to
- Question:
    - How does cell phone technology change market efficiency?
    - Key outcomes: price variation and wastage

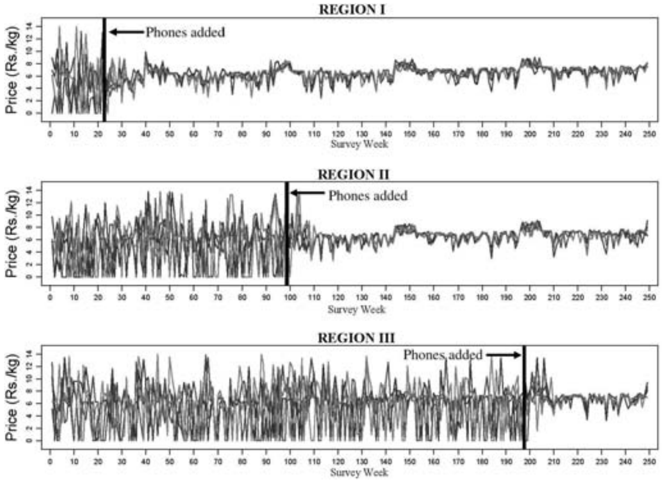
Setting


May 21, 2000
Kasaragod
Kanhangad
May 21, 2000
REGION III
(Kasaragod)
July 6, 1998
Kannur City
Thalassery
July 31, 1998
REGION II
(Kannur)
Kozhikode
Jan. 31, 1997
REGION I
(Kozhikode)

Technology adoption



Results



REGION I
Phones added
Price (Rs./kg)
Survey Week
REGION II
Phones added
Price (Rs./kg)
Survey Week
REGION III
Phones added
Price (Rs./kg)
Survey Week

Results

TABLE III
PRICE DISPERSION AND WASTE IN KERALA SARDINE MARKETS

	Period 0 (pre-phone)	Period 1 (region I adds phones)	Period 2 (region II adds phones)	Period 3 (region III adds phones)
Max–min spread (Rs/kg)				
Region I	7.60	1.86	1.32	1.22
	(0.50)	(0.22)	(0.10)	(0.44)
Region II	8.19	7.30	1.79	1.57
	(0.44)	(0.29)	(0.19)	(0.16)
Region III	8.24	7.27	7.60	2.56
	(0.47)	(0.27)	(0.25)	(0.34)
Coefficient of variation (percent)				
Region I	.68	.14	.08	.07
	(0.07)	(0.01)	(0.01)	(0.01)
Region II	.62	.55	.12	.08
	(0.04)	(0.04)	(0.01)	(0.01)
Region III	.69	.57	.54	.14
	(0.09)	(0.04)	(0.03)	(0.02)
Waste (percent)				
Region I	0.08	0.00	0.00	0.00
	(0.01)	(0.00)	(0.00)	(0.00)
Region II	0.05	0.04	0.00	0.00
	(0.01)	(0.01)	(0.00)	(0.00)
Region III	0.07	0.06	0.06	0.00
	(0.01)	(0.01)	(0.01)	(0.00)

High tech and development

- Jensen is a 'possibility result': shows an important example of how technology really can matter for poor people
- How much does it matter? Where else might it matter?
  - Political economy:
    - Freedom of information
    - Online procurement (Banerjee, Olken, and Pande in progress)
  - Monitoring:
    - Attendance of teachers, nurses, etc is a huge problem.
    - Cameras (Duflo, Hanna, and Ryan 2008)
    - Fingerprint readers in Indian schools
    - Compliance with drug regimes (e.g., using cell phones to monitor TB compliance)
  - Computer assisted learning
    - Banerjee et al (2007)

Appropriate technology

- Technology is not one-size-fits all
- Rather, technology is usually specific solutions to specific problems
- These problems are likely to differ in developing / developed countries.
  - Relative price of capital / labor is very different. So whether technology should augment capital or augment labor should be different. (Basu and Weil 1998)
  - Skill levels are different. So whether technology should augment high-skill labor or low-skill labor should also be different. (Acemoglu and Zilibotti 2001)
- Since technology is developed in proportion to market size, this usually means that technology is developed in rich countries and exported to poor countries, where it is suboptimal.
- Empirics of appropriate technology: best work calculates firm level overall productivity in different countries, and examines how patterns differ by high and low skill (Acemoglu and Zilibotti). But would be nice to show something even more direct.

Example: garbage collection

- In US:

Photograph of a garbage truck removed due to copyright restrictions.

Example: garbage collection

- In Indonesia:

Photograph of men manually hauling garbage removed due to copyright restrictions.

International technology adoption

- As discussed already, most technologies are developed in rich countries
- How do they get to poor countries?
- One view is that firm linkages are important – e.g, foreign direct investment, joint ventures etc.
- Empirics: once again, most empirical work investigates this by looking at changes in firm level productivity. Would be nice to see something more direct actually showing technology transfer.

Aitken and Harisson (1999)

- Examine whether FDI changes the productivity of manufacturing firms in Venezuela.
  - Also examine spillovers: once technology is transferred to a particular firm through FDI, what happens to other firms
- Fixed effects regression:

$$y_{ft} = \alpha_f + \alpha_t + \beta FDI_{ft} + \gamma X_{ft} + \varepsilon_{ft}$$

- Include interactions:

$$y_{fit} = \alpha_f + \alpha_t + \beta_1 FDI_{ft} + \beta_2 FDI_{it} + \beta_3 FDI_{fit} * FDI_{it} + \gamma X_{fit} + \varepsilon_{fit}$$

- Thoughts on this regression?
  - Why do foreign firms invest in particular firms?
  - Technology spillovers vs. product market effects? (i.e., if you lose business, your TFP may go down if you cannot easily adjust capital)

Aitken and Harisson (1999)

- Find positive own effects, negative spillovers, but only for small firms.

TABLE 3—IMPACT OF FOREIGN OWNERSHIP BY PLANT SIZE:
REGRESSING LOG OUTPUT AT THE PLANT LEVEL ON INPUTS AND THE SHARE OF FOREIGN OWNERSHIP AT THE PLANT LEVEL, THE SECTOR LEVEL, AND THE LOCAL LEVEL[a]

	Small plants (less than or equal to 49 employees)				Large plant (greater than 49 employees)			
	(1) OLS	(2) Within[b]	(3) OLS	(4) Within[b]	(5) OLS[c]	(6) Within[b,c]	(7) OLS[c]	(8) Within[b,c]
Foreign ownership in the plant (*Plant_DFI*)	0.104	0.100	0.167	0.182	0.121	−0.018	0.174	−0.123
	(0.052)	(0.055)	(0.065)	(0.084)	(0.031)	(0.049)	(0.036)	(0.073)
Foreign ownership in the sector and region (*Local_Sector_DFI*)	—	—	0.061	0.072	—	—	−0.020	0.196
			(0.035)	(0.058)			(0.032)	(0.218)
Plant_DFI * *Local_Sector_DFI*	—	—	−0.395	−0.359	—	—	−0.203	−0.285
			(0.138)	(0.170)			(0.080)	(0.247)
Foreign ownership in the sector over all regions) (*Sector_DFI*)	−0.349	−0.340	−0.366	−0.363	−0.127	−0.214	−0.128	−0.180
	(0.074)	(0.074)	(0.076)	(0.093)	(0.105)	(0.111)	(0.113)	(0.173)
Plant_DFI * *Sector_DFI*	1.184	0.046	1.475	0.559	0.351	0.411	0.590	1.033
	(0.595)	(0.564)	(0.584)	(0.837)	(0.205)	(0.279)	(0.225)	(0.372)
Number of observations	29,179	29,179	28,069	28,069	13,831	13,831	13,264	13,264
Number of plants	7,620	7,620	7,563	7,563	2,637	2,637	2,627	2,627
R^2	0.90	0.96	0.90	0.94	0.90	0.94	0.90	0.96

[a] Industry dummies included in all OLS specifications. All standard errors (denoted in parentheses) are corrected for heteroskedasticity. Unless otherwise specified, other independent variables (not reported here) include log materials, log skilled labor, log unskilled labor, and log capital stock. *Plant_DFI* is percentage of equity owned by foreigners. *Sector_DFI* is employment-weighted percentage of equity which is foreign owned at the four-digit ISIC level.
[b] Estimated by subtracting from each variable its plant-specific mean over all years.
[c] Regional controls include the real skilled wage and energy prices.

International technology adoption

- This is not the last word on this subject – would be nice to say something more direct.

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

14.771: Labor Lecture 1

Ben Olken

November 2008

Outline

- Today: how efficient are labor markets
  - Motivation: the surplus labor hypothesis
  - How well do labor markets work in developing countries?
    - Rural
    - Implications of having rural labor markets
    - Urban
  - Labor market regulation
- Next time: migration

Motivation

- Many people have observed high rates of "unemployment" or "underemployment" in rural areas
- For example, data from Walker and Ryan (1990 ICRISAT study)
  - Men – 19% (slack season 39%, peak season 12%)
  - Women – 23% (slack season 50%, peak season 11%)
  - But labor markets exist: 60-80% of labor use is hired

Surplus labor

- Lewis (1954), drawing on even older theories (e.g., Marx), argued that there was "surplus labor" in the countryside. He argued that "about 25%" of labor had zero marginal value.
  - Claim: you can move labor from countryside to cities without decreasing agricultural output
  - This would mean that either:
    - The marginal product of labor is zero because the agricultural production function is Leontief.
    - Labor supply is totally elastic at some reservation wage rate.

Shultz's (1964) test

- One of the earliest natural experiment studies
- Studies the 1917-1918 influenza epidemic, which killed 6% of the population and reduced the workforce by about 8%
- Idea: if there really was 25% surplus labor, then agricultural output would not fall!
- Empirics:
  - Compares output in 1919–1920 to 1916-1917, which had similar weather
  - Looks at whether provinces with greater influenza deaths had greater declines in output
  - Examines acres sown, since does not have direct data on output

Shultz's (1964) test

Deaths Attributed to Influenza Epidemic of 1918-19, and Predicted and Observed Effects on Agricultural Production for India and Major Provinces of India

Province and all India	*A measure of the distribution of deaths (per 100 population)*	*Adjusted distribution of deaths (per 100 population)*	*Predicted reduction in agricultural production (in percent)*	*Observed reduction in acreage sown to crops (in percent)*
(1)	*(2)*	*(3)*	*(4)*	*(5)*
Central Province and Berar	6.64	15.60	8.32	7.00
Bombay	5.49	12.90	6.88	2.10
Punjab	4.54	10.67	5.69	8.20
North West Frontier Province	4.36	10.25	5.47	7.00
United Province	4.34	10.20	5.44	6.60
Bihar-Orissa	2.05	4.82	2.57	+0.50
Assam	1.86	4.37	2.33	3.60
Madras	1.67	3.92	2.09	2.20
Burma	1.39	3.27	1.74	+4.00
Bengal	0.85	2.00	1.07	1.40
All British India	2.64	6.20	3.30	3.80

Figure by MIT OpenCourseWare.

Shultz's (1964) test

- Finds elasticity of output with respect to population of about 0.4, statistically significant with only 10 states!
- Does this rule out surplus labor? What assumptions would be required?
  - Land would need to be reallocated

Benjamin (1992)

- Question: how effecient are rural labor markets?
- Test of this: are production and consumption decisions 'separable'? This is the 'separation hypothesis'.
- Theoretical idea: with fully functioning efficient markets, households can freely buy or sell labor at wage w.
- Households therefore choose:
  - The labor input for their farms to maximize profits given wage w
  - The optimal labor/leisure tradeoff for the family given w
- With full ability to buy and sell labor at w there is no reason these two decisions should be related
- Empirical test: do household demographic characteristics (which should affect labor supply) affect labor demand for the family firm?
- Related to a milder view of "surplus labor": if there are labor market frictions, you may employ labor on your farm even if the marginal product is below the outside market wage. You'll do you more of this if you have more people available in your household.

Separation

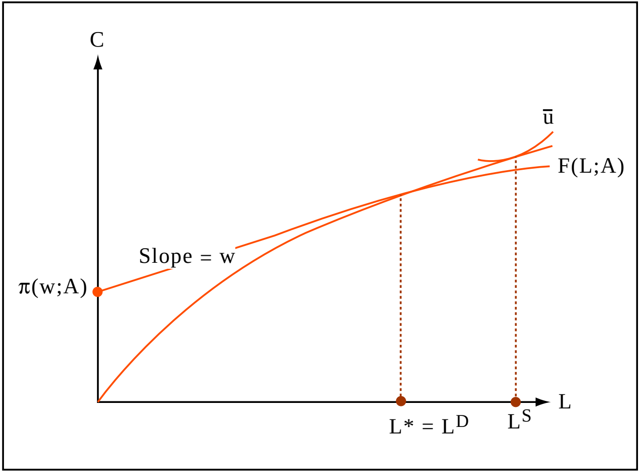


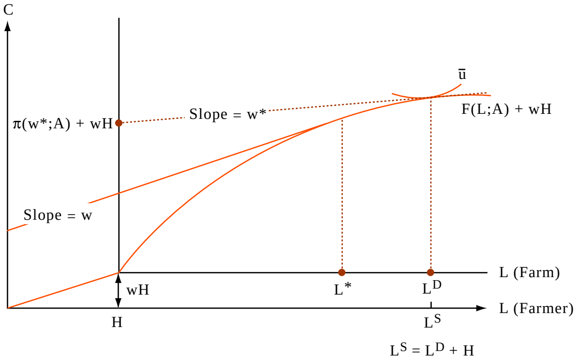
Figure by MIT OpenCourseWare.

Limits to separation

- When might separation not hold?
  - Minimum wage (implies maximum number of hours worked outside farm)
  - Imperfect labor markets (outside wage lower than inside wage)
  - Agency problems on land (efficiency of outside labor is lower)
  - Other market failures?

Examples of non-separation

- Suppose there is rationing (at H) in amount of off-farm work, because off farm wage is "too high". "Slack season"

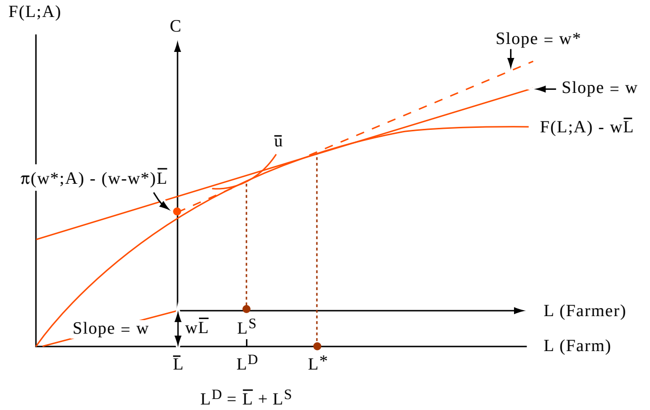


Case 1 constraint *H* on off-farm labor supply.

Figure by MIT OpenCourseWare.

Examples of non-separation

- Suppose there is rationing on hired labor at $\bar{L}$, because market wage is "too low." "Peak season"

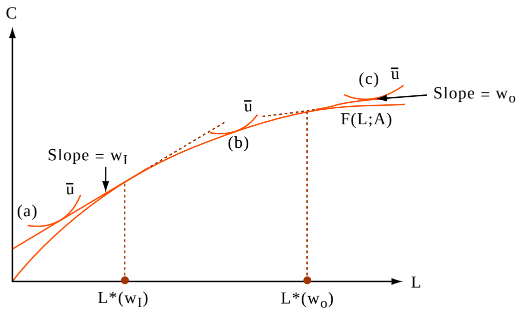


Case 2 constraint $\bar{L}$ on hired labor.

Figure by MIT OpenCourseWare.

Limits to separation

- Hired labor costs more than farmer's return to off-farm employment ($w_I > w_O$). E.g., agency problems in farm labor.



Case 3 $w_I > w_o$.

Figure by MIT OpenCourseWare.

- Note: empirics in case (a) depends on whether this is really a different in wage (in which case it looks like separation) or whether it is due to unobserved effective efficiency (in which case it doesn't look like separation)

Empirics

- Data:
  - 1980 SUSENAS from Java
- Estimation strategy:
  - Estimate labor demand, and see if it depends on demographics

$$\log L = \alpha + \beta \log w + \theta \log A + \delta_0 \log n + \sum_{i=1}^{D-1} \delta_i \frac{n_i}{n} + \varepsilon$$

Empirical issues: division bias

- Benjamin mentions a concern about "division bias." What is this?
- He calculates the wage by dividing the total wage bill by labor demand, i.e., $w = \frac{wages}{L}$
- Suppose L is measured with error, i.e., $L = L^* + \upsilon$
- You regress

$$L = \alpha + \beta \log w + \varepsilon$$

- Substituting for the measurement you get

$$L^* + \upsilon = \alpha + \beta \left(\log wages - \log L^* - \log \upsilon \right) + \varepsilon$$

- So now x is negatively correlated with the error term, which yields a downward bias of β
- Solution: instrument for wages with something uncorrelated with L^*, in his case, the wages of everyone else in the village

Empirical issues: simultaneity bias

- Wages are not exogenous – they are determined by the equilibrium of supply and demand.
- Regressing labor quantity on wage does not necessarily recover labor demand elasticity (the object of interest).
- For example, an aggregate labor demand shock, such as a positive productivity shock to agriculture, would increase labor demand and increase the wage, biasing the coefficient upward
- This is the classic argument for IV – we need instruments for labor supply that do not affect labor demand
- Benjamin uses population density to instrument for labor supply. Good instrument?

Empirical issues: endogeneity of household size

- Key question: under the null of separation between labor supply and labor demand ($\delta = 0$), can endogenous household size generate a false rejection of the null?
- Note what is not a problem:
  - Suppose separation does not hold, and household size expands to meet periods of peak labor demand.
  - Then we would find $\delta \neq 0$, because we'd find greater household size leads to greater labor demand.
  - That might be biased in the sense that we're not identifying the causal impact of exogenous household size on labor demand
  - BUT it would lead us to reject the null that $\delta = 0$, and it would do so precisely because $\delta \neq 0$.
- What could be a problem?
  - Family labor is measured more accurately than hired labor.
  - Omitted variables, i.e., better land quality $\rightarrow$ higher income $\rightarrow$ more kids and better land qualtiy $\rightarrow$ more labor demand
- His solution: district controls, cluster fixed effects, etc.

Results

- Most farmers have a mix of hired labor, family labor, and working outside farm: suggests fluid labor market

Cross Tabulation of Hiring-In and "Hiring-Out" for Rice Farmers

		Hired labor	*No hired labor*
Use family labor?	Yes	94.5	5.5
Wage employment last year?	Yes	46.0	2.0
	No	48.5	3.5
Nonagricultural employment last year?	Yes	82.3	4.9
	No	12.2	0.6
Work off-farm last week?	Yes	39.8	1.9
	No	54.7	3.6

Figure by MIT OpenCourseWare.

Results

Demand for Pre-Harvest Labor
Dependent Variable: Log Person Days Employed
(Standard Errors in Parentheses)
(*p* Values for *F* Tests)

	Parsimonious OLS	*Full OLS*	*Excluding children OLS*	*Within cluster*	*2SLS (meas. error)*	*2SLS (simultaneity)*	*2SLS (simultaneity and log h)*	*Means*
Intercept	4.780	2.085	2.255		2.343	2.657	2.623	
	(0.119)	(0.533)	(0.532)		(0.543)	(0.682)	(0.663)	
Log area harvested	0.680	0.680	0.682	0.696	0.686	0.757	0.742	-0.823
	(0.018)	(0.017)	(0.017)	(0.018)	(0.018)	(0.036)	(0.038)	
Log wage	-0.296	-0.274	-0.274		-0.315	-0.939	-0.894	0.912
	(0.027)	(0.026)	(0.026)		(0.040)	(0.252)	(0.231)	
Log pesticide price		0.139	0.139	-0.058	0.149	0.157	0.155	2.663
		(0.042)	(0.042)	(0.062)	(0.043)	(0.051)	(0.050)	
Log fertilizer price		0.407	0.409	0.401	0.367	0.409	0.405	4.301
		(0.111)	(0.111)	(0.107)	(0.117)	(0.135)	(0.132)	
Not irrigated		-0.147	-0.147	-0.070	-0.172	-0.156	-0.157	0.385
		(0.034)	(0.034)	(0.053)	(0.035)	(0.042)	(0.041)	
Log household size	0.043	0.078	0.068	0.052	0.097	0.032	0.039	1.471
	(0.045)	(0.046)	(0.044)	(0.045)	(0.049)	(0.059)	(0.057)	
Prime male fraction	-0.058	0.079	-0.075	0.127	0.094	0.015	0.023	0.256
	(0.108)	(0.105)	(0.127)	(0.100)	(0.109)	(0.130)	(0.127)	
Prime female fraction	-0.163	0.019	-0.133	0.106	0.067	0.004	0.004	0.272
	(0.128)	(0.128)	(0.109)	(0.116)	(0.131)	(0.156)	(0.152)	
Elderly male fraction	0.043	0.279	0.129	0.194	0.280	0.208	0.220	0.055
	(0.145)	(0.187)	(0.167)	(0.171)	(0.198)	(0.230)	(0.224)	
Elderly female fraction	-0.076	0.166		0.085	0.129	0.051	0.053	0.051
	(0.151)	(0.163)		(0.150)	(0.173)	(0.203)	(0.198)	
Age of head		0.013	0.014	0.009	0.010	0.012	0.012	45.588
		(0.007)	(0.007)	(0.006)	(0.007)	(0.009)	(0.009)	
Age squared		-0.00015	-0.00010	-0.0001	-0.0001	-0.0001	-0.0001	2241.534
		(0.00008)	(0.00007)	(0.00007)	(0.00008)	(0.00009)	(0.00009)	
F Education of head		2.91	3.06	1.83	2.95	1.38	1.45	
		(0.008)	(0.006)	(0.089)	(0.007)	(0.22)	(0.189)	
F Kabupaten soil		7.92	7.95		8.69	6.76	7.06	
		(0.0001)	(0.0001)		(0.0001)	(0.0001)	(0.0001)	
F Kabupaten climate		13.05	13.34		7.23	3.56	3.86	
		(0.0001)	(0.0001)		(0.0001)	(0.014)	(0.009)	
Sugar regency		0.135	.135		0.115	0.110	0.111	0.447
		(0.033)	(0.033)		(0.035)	(0.041)	(0.040)	
F Dems	1.19	1.03	1.55	0.53	1.03	0.237	0.279	
	(0.311)	(0.399)	(0.18)	(0.756)	(0.396)	(0.948)	(0.924)	
R-Squared	0.525	0.591	.591	0.872	0.6017	0.473	0.408	
Wu-Hausman					3.67	(7.05)	(7.75)	

Figure by MIT OpenCourseWare.

Results

Implied Demographic Elasticities from Table IV
(Standard Errors in Parentheses)

Specification:	*Elasticity of labor demand with respect to additional household members:*						
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Type of member:							
Prime age male	0.012 (0.024)	0.028 (0.025)	0.010 (0.018)	0.027 (0.024)	0.032 (0.026)	0.007 (0.031)	0.010 (0.030)
Prime age female	-0.016 (0.025)	0.013 (0.027)	-0.004 (0.019)	0.022 (0.025)	0.027 (0.028)	0.005 (0.003)	0.006 (0.032)
Elderly male	0.008 (0.005)	0.017 (0.006)	0.013 (0.005)	0.010 (0.005)	0.017 (0.006)	0.012 (0.006)	0.013 (0.007)
Elderly female	0.001 (0.005)	0.010 (0.005)	0.006 (0.004)	0.003 (0.005)	0.008 (0.005)	0.003 (0.006)	0.004 (0.006)
Child (< 15 yrs)	0.038 (0.018)	0.011 (0.017)		-0.007 (0.016)	0.012 (0.018)	0.005 (0.020)	0.006 (0.021)

Specifications: (1) Parsimonious OLS. (2) OLS with full set of control variables. (3) OLS with full set of control variables, but children under 15 yrs. excluded from household size. (4) Within cluster estimation. (5) 2 SLS for correction of measurement error of wage. (6) 2SLS for correction for potential simultaneity of wage. (7) 2SLS for correction for potential simultaneity of wage and adjustment of area harvested.

Figure by MIT OpenCourseWare.

Results

- Finds demography does not affect labor demand
- Interprets this as evidence that labor demand and labor supply are 'separable' – i.e., rural labor markets actually work pretty well.
- Do you find this persuasive?

Jayachandran (2006)

- Benjamin's paper suggests that rural labor markets exist and are relatively active
- This implies that rural households' earnings depend not just on their own farm's productivity, but are also determined by the aggregate wage rate
- Jayachandran's idea:
    - The rural wage will be more inelastic if workers are unable to smooth shocks. In particular it will be more inelastic if there is:
        - Less access to credit
        - Lower ability to migrate
    - Inelastic wages imply larger impacts of productivity shocks on rural welfare.
    - They also imply a pecuniary externality – it is not just your own ability to smooth that affects your ability to cope with shocks, but the ability of everyone else around to smooth also affects your welfare.

Empirical idea

- Empirical goal: estimate labor supply elasticity
- Therefore we need an instrument for labor demand
- Jayachandran uses rainfall shocks as instrument for labor demand:
  - $Rainshock = 1$ if above 80th percentile of rain, 0 if between 20th and 80th, and -1 if below 20th percentile
- Estimating equation:

$$w_{jt} = \beta_1 A_{jt} + \beta_2 S_{jt} + \beta_3 S_{jt} \times A_{jt} + \beta_4 X_{jt} + \beta_5 X_{jt} \times A_{jt} + \delta_t + \alpha_j + \varepsilon_{jt}$$

  where key coefficients of interest are β_3
- Instruments for A_{jt}, $S_{jt} \times A_{jt}$, $X_{jt} \times A_{jt}$ with $Rainshock_{jt}$, $S_{jt} \times Rainshock_{jt}$, $X_{jt} \times Rainshock_{jt}$

First stage

	Dependent variable		
	Log crop yield: OLS (1st stage)	*Log agricultural wage: OLS*	*Log agricultural wage: Instrumental variables*
	(1)	*(2)*	*(3)*
Rainshock	.070*** (.007)		
Rainshock x %Agrarian	.003 (.005)		
Log crop yield		.035*** (.012)	.167** (.084)
Log crop yield x %Agrarian			-.009 (.039)
Observations	8,222	8,222	8,222
District and year fixed effects?	Yes	Yes	Yes

Figure by MIT OpenCourseWare.

Results

Banking and the Elasticity of the Wage
Dependent Variable: Log Agricultural Wage, 1956-87

	Measure of banking		
	Bank deposits per capita	*Bank credit per capita*	*Bank branches per capita*
	(1)	*(2)*	*(3)*
Log crop yield	.162** (.083)	.158* (.083)	.138* (.082)
Banking			-.049** (.021)
Log crop yield x banking	-.091** (.036)	-.075* (.044)	-.033* (.019)
Observations	7,678	7,614	8,080
District and year fixed effects?	Yes	Yes	Yes

Figure by MIT OpenCourseWare.

Results

Access to Neighboring Areas and the Elasticity of the Wage
Dependent Variable: Log Agricultural Wage, 1956-87

	Measure of access to neighboring areas			
	Road density (km/km^2)	*Bus service (% of villages)*	*Railway (% of villages)*	*Closeness to city (km^{-1})*
	(1)	*(2)*	*(3)*	*(4)*
Log crop yield	.133* (.080)	.147* (.076)	.162** (.082)	.171** (.084)
Access	-.026 (.020)			
Log crop yield x Access	-.111 (.083)	-.095* (.046)	-.098* (.051)	-.050 (.039)
Observations	7,965	7,838	7,838	8,222
District and year fixed effects?	Yes	Yes	Yes	Yes

Figure by MIT OpenCourseWare.

Results

Poverty, Land Inequality, and the Elasticity of the Wage
Dependent Variable: Log Agricultural Wage, 1956-87

	District trait			
	Poverty		*Land Inequality*	
	Average expenditure	*Poverty head count*	*%Landless*	*Gini coefficient*
	(1)	*(2)*	*(3)*	*(4)*
Log crop yield	.183** (.090)	.181** (.091)	.121 (.084)	.186** (.091)
District trait			-.059** (.026)	
Log crop yield x district trait	-.034 (.028)	-.002 (.045)	-.157*** (.056)	-.005 (.048)
Observations	7,934	7,934	8,222	7,711
District and year fixed effects?	Yes	Yes	Yes	Yes

Figure by MIT OpenCourseWare.

More on flexible labor markets and shocks

- Jayachandran shows using micro-data that agricultural wages respond to productivity shocks
- Do they respond enough for markets to clear? Is this true for an entire economy?
- Smith et al (2002) examine the case of the Asian financial crisis in Indonesia
  - This is a massive shock: currency drops from Rp.2,500/$ to as low as Rp.14,000, real GDP declines by 13% in 1998
  - Question: how much of this absorbed by unemployment, and how much by changes in real wages?

Results

Provinces:	*% Change between 1986 and 1997*			*% Change between 1997 and 1998*		
	All	*IFLS*	*IFLS2+*	*All*	*IFLS*	*IFLS2+*
	(1)	*(2)*	*(3)*	*(1)*	*(2)*	*(3)*
Males						
Wage	40.3	41.5	42.2	-37.8	-38.1	-36.0
Urban sector	36.2	37.1	37.3	-40.6	-40.9	-38.6
Rural sector	38.2	38.7	39.4	-35.6	-35.9	-33.7
% Working	-0.5	-0.7	-0.3	-0.9	-1.1	-1.7
In wage sector	7.1	7.1	6.4	-3.0	-2.7	-2.5
Self-employed	-3.3	-3.8	-3.5	1.5	1.0	1.3
Unpaid family	-4.1	-4.0	-4.6	0.6	0.6	0.6
Females						
Wage	62.1	63.9	67.7	-37.9	-38.3	-39.2
Urban sector	62.5	64.8	68.7	-41.3	-41.8	-43.3
Rural sector	52.6	52.6	52.7	-33.9	-33.8	-32.7
% Working	-2.3	-0.9	-1.3	1.0	1.3	1.0
In wage sector	3.7	3.8	3.8	-0.3	-0.2	-0.2
Self-employed	1.1	0.6	1.2	0.5	0.4	0.5
Unpaid family	-6.4	-5.5	-6.1	0.8	1.1	0.8

Figure by MIT OpenCourseWare.

- Note: substantial inflation may have allowed more wage flexibility than normal. Still, pretty interesting.

Besley and Burgess (2004): Restrictions on labor market flexibility

- What happens when governments impose labor regulations?
  - Labor regulations seek to provide better working conditions, etc.
  - But may reduce the returns for firms
- Test from India, where labor regulation occurs at state level
- Code each state ammendment to industrial law as pro-worker (1), neutral (0), or pro-firm (-1)
- Run a differences-in-differences regression

$$y_{st} = \alpha_s + \beta_t + \mu r_{st-1} + \varepsilon_{st}$$

Do regulations matter?

TABLE II
LABOR REGULATION AND INDUSTRIAL DISPUTES IN INDIA: 1958–1992

	(1)	(2)	(3)	(4)
	Workdays lost to strikes per worker	Workdays lost to strikes per worker	Workdays lost to lockouts per worker	Workdays lost to lockouts per worker
Method	OLS	OLS	OLS	OLS
Labor regulation	2.564**	1.732*	2.108**	0.965***
[t − 1]	(2.55)	(1.87)	(2.32)	(3.57)
State effects	YES	YES	YES	YES
Year effects	YES	YES	YES	YES
State time trends	NO	YES	NO	YES
Adjusted R^2	0.08	0.07	0.14	0.15
Observations	547	547	514	514

Impact on manufacturing

TABLE V

Labor Regulation and Employment, Investment, and Productivity in Registere

	(1)	(2)	(3)
	Log registered manufacturing employment	Log daily employment in registered manufacturing	Log earnings per worker in registered manufacturing
Method	OLS	OLS	OLS
Labor regulation	−0.072*	−0.285***	0.008
[t − 1]	(1.70)	(3.48)	(0.09)
Log development expenditure	0.076	0.327*	0.207
per capita	(0.64)	(1.82)	(1.52)
Log installed electricity	0.073	0.111	0.019
capacity per capita	(1.34)	(1.51)	(0.34)
Log state population	−0.099	2.122	1.116
	(0.09)	(1.14)	(0.93)
Congress majority	0.008	−0.009	−0.037*
	(0.61)	(0.39)	(1.66)
Hard left majority	−0.028	−0.124***	0.0004
	(1.43)	(3.93)	(0.01)

Aggregate impact

	(1)	(2)	(3)	(4)	(5)
	Log state output per capita	Log state agricultural output per capita	Log state nonagricultural output per capita	Log state construction output per capita	Log total manufacturing output per capita
Method	OLS	OLS	OLS	OLS	OLS
Labor regulation [$t-1$]	−0.002	0.019*	−0.034*	−0.019	−0.073**
	(0.14)	(1.81)	(1.69)	(0.29)	(2.05)
State effects	YES	YES	YES	YES	YES
Year effects	YES	YES	YES	YES	YES
Adjusted R^2	0.93	0.84	0.95	0.76	0.93
Observations	509	509	509	509	509

Conclusions

- Evidence that rural labor markets work pretty well
  - This implies pecuniary externalities from other people's smoothing
- Urban labor regulations don't seem to help workers, but do reduce manufacturing output.

14.771 Development Economics: Microeconomic issues and Policy Models
Fall 2008

14.771: Labor Lecture 2

Ben Olken

November 2008

Overview

- Migration
  - International migration and wage differentials
  - Networks among migrants

Migration across countries

- Wage differences across countries are vast, and suggest massive inefficiencies
  - ILO (1995) reports that the wage of a construction carpenter in India was $42 per month ($250 or so at PPP), vs. $2200 in the US.
  - How much of this is difference in skill and how much a difference in the role of other inputs and/or TFP?
  - To the extent it is other inputs and/or TFP, there is a first order gain in efficiency when you move the carpenter to the US.
- However there is also a brain-drain/brain-gain:
  - Complementary factors capture part of the gain in the destination countries. Close substitutes lose.
  - Complementary factors lose in the sending countries.
  - So wage differences do not exactly predict the magnitude of total general equilibrium changes we would expect if we allowed full migration

The skill price and migration

- A simple way to think about these issues is to think of a 'skill price', i.e.,
  - Person j in country i earns a wage $W_{ij} = w_i x_{ij}$ where w_i is the skill price in country i and x_{ij} is j's amount of skill.
  - People migrate from i to u if
  $$w_u x_{ij} - C_i \geq w_i x_{ij}.$$
  where C_i is the cost of migration from that country to country u (for U.S.)
  - So migration occurs if
  $$x_{ij} \geq \frac{C_i}{w_u - w_i}$$
- This simple framework has several implications
  1. Only the most skilled people from a given country will migrate.
  2. The higher the skill price of a sending country (w_i), the higher is the skills of the migrants from that country.
  3. The harder it is to get to the US from that particular country (C_i), the higher is the skill level of those who migrate from that country.

Estimating the skill price in theory

- We would like to estimate these skill prices to test models of cross-country migration
- How can one estimate the skill price?
- Assume that

$$x_{ij} = \exp[\beta_i(Q_i)S_{ij} + \sum_k \gamma_{jk} I_{ijk} + \mu_{ij}]$$

is the skill production function, where S_{ij} is the years of schooling, $\beta_i(Q_i)$ is the return on schooling as a function of quality of schooling in that country, and I_{ijk} is a set of other human capital attributes.
- Wages in a given country are determined by

$$W_{ij} = w_i x_{ij}$$

and hence

$$\log W_{ij} = \log w_i + \beta_i(Q_i)S_{ij} + \sum_k \gamma_{ijk} I_{jk} + \mu_{ij}.$$

- The intercept of this equation is the skill price.

Estimating the skill price in practice

- The problem is that to estimate this equation, one needs comparable micro data on wages, schooling, and other human capital variables for a large sample of countries.
- This does not appear to exist.
- Rosenzweig (2007) proposes an alternative – examining migrants who are in the US.
- He (and others) collected a dataset called the "New Immigrant Survey", which provides comparable data on immigrants from 140 countries on the last job they had before they came to the U.S.
- What are the pitfalls of this approach?
    - Selective sample into migration: model implies immigrants positively selected on unobservables. Why is this a problem?
    - Selective sample of countries: only those with sufficient immigrants in U.S. Why is this a problem?

Empirical note: the Heckman Selection Correction

- Selective sample into migration implies that w_i is biased upwards, because you are more likely to be in the sample if you have high unobservables.
- Rozensweig deals with this using a Heckman selection model:
  - Our migration model from earlier implies that migration is a function of determinants of country skill price (GDP, quantity of skills) and migration costs (distance):

    $$\Pr(migrate = 1) = \Pr\left(X'\gamma + \upsilon > 0\right)$$

    where υ is Normal and has correlation ρ with μ (error from wage equation).
  - The selection issue is we estimate the wage equation conditional on migrating, i.e., conditional on $X'\gamma + \upsilon > 0$.
  - Since μ and υ are correlated, this implies that

    $$\log W_{ij} = \log w_i + \beta_i(Q_i)S_{ij} + \sum_k \gamma_{ijk} I_{jk} + E\left(\mu_{ij} \mid X'\gamma + \upsilon > 0\right).$$

Estimation

- If both υ and μ are normal, then the conditional expectation is given by the Mills ratio:

$$E\left(\mu_{ij} \mid X'\gamma + \upsilon > 0\right) = \rho \frac{\phi\left(X'\gamma\right)}{\Phi\left(X'\gamma\right)}$$

- So we first estimate migration as a function of gdp, schooling levels, and migration costs:

$$\Pr\left(migrate = 1\right) = \Pr\left(\alpha + \beta_1 Y + \beta_2 S + \beta_3 C + \upsilon > 0\right)$$

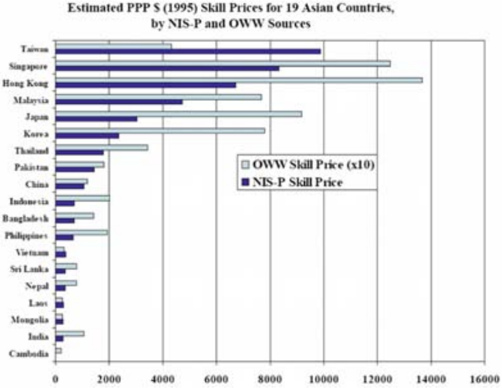
- And then use the resulting coefficients to estimate:

$$\log W_{ij} = \log w_i + \beta_i(Q_i) S_{ij} + \sum_k \gamma_{ijk} I_{jk} + \rho \frac{\phi\left(\alpha + \beta_1 Y + \beta_2 S + \beta_3 C\right)}{\Phi\left(\alpha + \beta_1 Y + \beta_2 S + \beta_3 C\right)}.$$

- Note:
  - If the same variables are in the main equation and the selection equation, you are entirely identified off function form.
  - More generally functional form matters a lot, so we prefer not to rely on these types of models.

Estimates of skill prices

- Skill prices differ significantly across countries
  - Skill price in S. Korea is 3.5 to 5.5 times that in Bangladesh! Taiwan is 20 times Bangladesh!



Courtesy of Mark Rosenzweig. Used with permission.

Determinants of skill price

- First estimate migration equation (do you migrate to US).
  - More migrants if your GDP is lower, if you are closer to the US, if there is more schooling in your country (which lowers skill price)

Determinants of the Probability of Immigration (NIS)

Origin-country variable		
PPPS GDP per worker ($x10^{-3}$)	-.0129 (3.44)	-.0126 (3.56)
Average adult schooling attainment	.0463 (2.64)	.0576 (3.80)
Distance to the United States (miles$x10^{-3}$)	-.0901 (5.60)	-.0985 (7.62)
US military base in country	.205 (2.46)	.177 (2.46)
Any ranked universities	-.225 (1.08)	-.214 (1.28)
Average rank of ranked universities	.000406 (0.25)	.000463 (0.34)
Teacher/pupil ratio in seconday schools	--	.00810 (1.59)
Teacher/pupil ratio in primary schools	--	.00368 (1.66)

Figure by MIT OpenCourseWare.

Determinants of skill price

- Estimate skill price equation using migration equation to correct for selection bias
  - Note that the variables are the same except distance to US – so distance to US (and functional form) is being used to identify selection
  - GDP increases skill price; more skilled labor decreases skill price

Estimates of the Determinants of the Country Log Skill Price

Sample	*US immigrant home wages (NIS-P)*	
Variable/Estimation procedure	*GLS*	*GLS-SC*
Country characteristics:		
Log GDP per worker	1.41 (5.01)	1.35 (5.21)
Log mean schooling	-1.77 (3.18)	-1.97 (3.23)
Log teacher-pupil ratio, primary schools	-1.90 (3.68)	-2.17 (3.80)
Log teacher-pupil ratio, secondary schools	1.44 (2.51)	1.36 (2.56)
Immigrant skill characteristics:		
Schooling	.0683 (3.50)	.0745 (3.79)
Age	.0428 (4.32)	.0436 (4.50)
Mills ratio	-	.800 (1.46)

Figure by MIT OpenCourseWare.

Determinants of skill level of migrants

- More educated migrants from higher skill price countries, farther countries

Determinants of the Log Average Schooling Attainment of New Adult Immigrants, by Type (NIS)

Origin-country variable/immigrant type	*Employment visa principals*	*EVPS's + spouses of citizens*	*All*
Log skill price	.499 (2.83)	.254 (1.98)	.139 (0.84)
Log real GDP per adult-equivalent	-.108 (1.60)	-.0116 (0.27)	.0557 (0.85)
Log distance of country to the United states	.0377 (4.43)	.0356 (4.53)	.0414 (3.89)
US military base in country	-.0220 (0.41)	-.0449 (1.39)	-.0127 (0.30)
English an official country language	.115 (1.62)	.0812 (2.84)	.120 (3.91)
Any ranked universities	1.18 (2.80)	.492 (1.78)	.115 (0.32)
Average rank of ranked universities	-.0110 (2.86)	-.00496 (1.77)	-.00162 (0.46)
Log teacher/pupil ratio in secondary schools	.00480 (0.15)	-.0230 (0.79)	.0380 (1.00)
Log teacher/pupil ratio in primary schools	-.00985 (0.08)	.0452 (1.10)	.0249 (0.40)

Figure by MIT OpenCourseWare.

So...

- Suggests that basic simple model of migration is explaining number and selection of migrants
- But data issues suggest this is far from the end of the story...

Networks - Munshi (2003)

- The simple model above assumed that the skill price was constant, and available to everyone – all you need to do is show up and you get the new wage
- In practice, labor markets are much more complicated, and people use networks to integrate themselves into economy in the new location.
- Some quotes from Munshi's paper:
    - "Leonardo shared an apartment with seven other friends, all paisanos from Sinaloa. Seven of the eight friends worked as gardeners. The first two friends had been in the area for five years, and provided referrals for employers for each of the subsequent migrants, the last of whom migrated two years earlier."
    - "Over 70 percent of the undocumented Mexicans, and a slightly higher proportion of the Central Americans, that Chavez interviewed in 1986 found work through referrals from friends and relatives."

Question

- Setting: migrants from Mexico to the US
- Question: what role do networks play in integrating migrants into the workforce?
- Empirical problem: If we observe that people from a village always go to the same place, it could be because they have common skills and there is a demand for those skills in that location.
- Munshi's solution:
  - Use lagged rainfall shocks at the origin location to instrument for the size of the network
  - Use individual fixed effects for migrants to control for selection in ability
- Data from the Mexican Migration Project

First stage: rainfall and migrants in US

- More new migrants ($<=$ 3 years) when recent rainfall low. More established migrants ($>$ 3 years) when old rainfall is low.

Dependent variable:	Reduced-form					First-stage	
	Employment at the destination				Employment at the origin	New migrants	Established migrants
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Recent-past rainfall	−0.028	−0.049	−0.023	−0.047	0.085	−0.091	0.005
	(0.027)	(0.035)	(0.072)	(0.040)	(0.018)	(0.037)	(0.020)
Distant-past rainfall	−0.125	−0.092	−0.226	−0.129	0.046	0.103	−0.106
	(0.035)	(0.027)	(0.108)	(0.044)	(0.021)	(0.033)	(0.023)
Individual fixed effects	Yes	Yes	Yes	No	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.705	0.705	0.647	0.038	0.812	0.768	0.940
Q statistic	0.041	0.041	0.036	0.660	2.813	0.010	0.316
Number of observations	4546	4546	1732	4546	41,120	4546	4546

Reduced form: rainfall and employment

- Recent rainfall has no effect on employment in US. Lagged rainfall has positive effect on US employment.
  - Also true when looks only at migrants who have arrived in past 1 or 2 years (column 3) though not much variation left after individual FE removed

Dependent variable:	Reduced-form					First-stage	
	Employment at the destination				Employment at the origin	New migrants	Established migrants
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Recent-past rainfall	−0.028	−0.049	−0.023	−0.047	0.085	−0.091	0.005
	(0.027)	(0.035)	(0.072)	(0.040)	(0.018)	(0.037)	(0.020)
Distant-past rainfall	−0.125	−0.092	−0.226	−0.129	0.046	0.103	−0.106
	(0.035)	(0.027)	(0.108)	(0.044)	(0.021)	(0.033)	(0.023)
Individual fixed effects	Yes	Yes	Yes	No	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.705	0.705	0.647	0.038	0.812	0.768	0.940
Q statistic	0.041	0.041	0.036	0.660	2.813	0.010	0.316
Number of observations	4546	4546	1732	4546	41,120	4546	4546

IV: employment

- Large effects of network size on probability of employment

					Employ
	OLS				
	Basic specifications				
Dependent variable:	(1)	(2)	(3)	(4)	(5)
New migrants	−0.032	0.397	0.522	0.093	0.626
	(0.070)	(0.315)	(0.376)	(0.537)	(0.501)
Established migrants	0.670	1.554	1.474	2.073	1.745
	(0.154)	(0.551)	(0.545)	(1.069)	(0.661)
Individual fixed effects	Yes	Yes	Yes	Yes	No
Year dummies	Yes	Yes	Yes	Yes	Yes
R^2	0.707	—	—	—	—
Q statistic	0.042	0.041	0.041	0.036	0.660
Number of observations	4546	4546	4546	1732	4546

IV: occupation

- Large effects of network size on probability of employment

Dependent variable:	Occupation at the destination			Occupation at the origin	Occupatio	
	Reduced-form				OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Recent-past rainfall	−0.125	−0.026	−0.025	0.011	—	—
	(0.077)	(0.217)	(0.096)	(0.015)		
Distant-past rainfall	−0.223	−0.479	−0.185	−0.027	—	—
	(0.086)	(0.258)	(0.098)	(0.019)		
New migrants	—	—	—	—	0.398	1.592
					(0.231)	(0.928)
Established migrants	—	—	—	—	0.256	3.585
					(0.247)	(1.339)
Individual fixed effects	Yes	Yes	No	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.825	0.734	0.143	0.898	0.825	—
Q statistic	0.319	0.007	1.605	1.886	0.322	0.319
Number of observations	4240	1588	4240	30,917	4240	4240

Limits to migration

- For international migration, limits are easy to understand: there are legal limits
  - Though who these limits benefit or hurt is an important question in the US labor literature – see Borjas, Card, Cortes, etc.
- For domestic migration, though, legal limits are usually not present
- But there may be other important limitations.
- Credit
  - Banerjee-Newman (1998) suggest that People in rural areas have access to informal insurance mechanisms, but when you migrate to an urban area, you lose this. This implies that migration will occur at the low end (where you have no collateral at all, so can't borrow either way) or at the high end (where you don't need informal insurance)
  - Munshi and Rosenzweig (2005) test these ideas in rural India
- Networks (same argument as Munshi applies domestically as well)
- Land market imperfections
- Behavioral issues

14.771 Development Economics: Microeconomic Issues and Policy Models
Fall 2008

Credit Access and the Poor

Abhijit Banerjee

Department of Economics, M.I.T.

1 The neo-classical model of the capital market

- Everyone faces the same interest rate, adjusted for risk. i.e. if there is a d% risk of default then $(1-d)r$ (where r is the gross interest rate) is a constant.

- The interest rate paid to depositors is equal to $(1-d)r$ less some small change for the cost of operating a bank.

- The expected marginal product of capital should be equated to $(1-d)r$.

2 Credit Markets: some facts

1. **Sizeable gap between lending rates and deposit rates within the same sub-economy:**

   Ghatak (1976) reports data on interest rates paid by cultivators in India from the *All India Rural Credit Survey* for the 1951-2 to 1961-2 period: The average rate varies between a maximum of 18% (in 1959-60) and a minimum of about 15% (in 1961-62). Around 25% of the borrowing reported in these surveys were zero-interest loans, usually from family members or friends. If these were left out, the average rates in these surveys would be above 20%. In comparison, Ghatak reports that the bond rate in this period was around 3% and the bank deposit rate was probably about the same.

Timberg and Aiyar (1984) report data on indigenous style bankers in India, based on surveys that they carried out. They report the gap between the average rate charged to borrowers and the average rate to depositors by Finance Companies was 16.5%. The same gap for financiers from the Shikarpuri community was 16.5%, 12% for financiers from the Gujerati community, 15.5% for the Chettiars, 11.5% for the Rastogis, etc.

The "Summary Report on Informal Credit Markets in India" (Dasgupta, 1989) reports that for the rural sector (the data is based on surveys of 6 villages in Kerala and Tamil Nadu), the average interest rate charged by professional money-lenders (who provide 45.61% of the credit) in these surveys is about 52%, while the average deposit rate is not reported, the maximum from all the case studies is 24% and the maximum in four out of the eight case studies is no more than 14%.

For the urban sector, the data is based on various case surveys of specific classes of informal lenders: For Finance Corporations they report that the maximum deposit rate for loans of less than a year is 12% while the minimum lending rate is 48%. For hire-purchase companies in Delhi, the deposit rate was 14% and the lending rate was at least 28%. For auto-financiers in Namakkal, the gap between the deposit rate and the lending rate was 19%. For handloom financiers in Bangalore and Karur, the gap between the deposit rate and the lowest lending rate was 26%.

Aleem (1990) reports data from a study of professional moneylenders that he carried out in a semi-urban setting in Pakistan in 1980-1981. The average interest rate charged by these lenders is 78.5%. The opportunity cost of capital to these money-lenders was 32.5%.

2. **Extreme variability in the interest rate within the same sub-economy:**

Timberg and Aiyar (1984) report that the rates for Shikarpuri financiers varied between 21% and 37% on loans to members of local Shikarpuri associations and between 21% and 120% on loans to non-members (25% of the loans were to non-members and another 50% were loans through brokers). On the other hand, the Gujerati bankers charged rates of no more than 18%. Moreover, the rates faced by established commodity traders in the Calcutta and Bombay markets were never above 18% and could be as low as 9%.

The "Summary Report on Informal Credit Markets in India" (Dasgupta, 1989) reports that Finance Corporations offer advances for a year or less at rates between 48% per year and the utterly astronomical rate of 5% per day. The rates on loans of more than a year varied between 24% and 48%. Hire-purchase contracts offer rates between 28% to 41% per year. Handloom Financiers charge rates between 44% and 68%. Yet the Shroffs of Western India offer loans at less than 21% and Chit Fund members can borrow at less than 25%.

The same report tells us that among rural lenders, the average rate for professional money-lenders (who in this sample give about 75% of the commercial informal loans) was 51.86%, whereas the rates for the agricultural money-lenders (farmers who also lend money) who supply the rest was 29.45%. Within the category of professional money-lenders, about half the loans were at rates of 60% or more but another 40% or so had rates below 36%.

The study by Aleem (1990) reports that the standard deviation of the interest rate was 38.14% compared to an average lending rate of 78.5%. In other words, an interest rate of 2% and an interest rate of 150% are both within two standard deviations of the mean.

Swaminathan (1991) reports on a survey of two villages in South India that she carried out: The average rate of interest in one village varied between 14.8% for loans collateralized by immovable assets (land, etc.) and 60% for loans backed by moveable assets. The corresponding rates in the other village were 21% and 70.6 %. Even among loans collateralized by the same asset—gold—the average rate in one village was 21.8% but it went up to 58.8% when the loans were to landless laborers.

Ghate (1992) reports on a number of case studies from all over Asia: The case study from Thailand found that interest rates were 2-3% per month in the Central Plain but 5-7% in the north and north-east (note that 5 and 7 are very different).

Gill and Singh (1997) report on a survey of 6 Punjab villages they carried out. The mean interest rate for loans up to Rs 10,000 is 35.81% for landowning households in their sample, but 80.57% for landless laborers.

Fafchamps' (2000) study of informal trade credit in Kenya and Zimbabwe reports an average monthly interest rate of 2.5% (corresponding to annualized rate of 34%) but also notes that this is the rate for the dominant trading group (Indians in Kenya, whites in Zimbabwe) is 2.5% month while the blacks pay 5% per month in both places.

Irfan et al. (1999), mentioned above, report that interest rates charged by professional money-lenders vary between 48% and 120%.

3. **Low levels of default:**

Timberg and Aiyar (1984) report that average default losses for the informal lenders they studied ranges between 0.5% and 1.5% of working funds.

The "Summary Report on Informal Credit Markets in India" (Dasgupta, 1989) attempts to decompose the observed interest rates into their various components, and finds that the default costs explain 14 per cent (not 14 percentage points!) of the total interest costs for the Shroffs, around 7% for auto-financiers in Namakkal and handloom financiers in Bangalore and Karur, 4% for Finance Companies, 3% for hire-purchase companies and essentially nothing for the Nidhis.

The same study reports that in four case studies of money-lenders in rural India they found default rates explained about 23% of the observed interest rate.

The study by Aleem gives default rates for each individual lender. The median default rate is between 1.5 and 2% and the maximum of 10%.

4. **There seems to be ex ante competition in the markets**

   Large numbers of lenders in any sub-market

- Aleem (1989) shows that lenders do not earn excess profits on average

  The "Summary Report on Informal Credit Markets in India" (Dasgupta, 1989) claims that only a small part of the interest rate is explained by profits.

  Ghate (1992) echoes the same conclusion.

5. **Production and trade finance are the main reasons given for borrowing, even in cases where the rate of interest is relatively high:**

   Ghatak (1976) concludes on the basis of his study that "the existing belief about the unproductive use of loans by Indian cultivators ... has not been substantiated."

   Timberg and Aiyar (1984) report that for Shikarpuri bankers (who charge 31.5% on average, and as much as 120% on occasion), at least 75% of the money goes to finance trade and, to lesser extent, industry.

The “Summary Report on Informal Credit Markets in India” (Dasgupta, 1989), reports that several of the categories of lenders that have been already mentioned, such as hire-purchase financiers (interest rates between 28%-41%), handloom financiers (44%-68%), Shroffs (18%-21%) and Finance Corporations (24%-48% for longer term loans and more than 48% on loans of less than a year) focus almost exclusively on financing trade and industry, and even for Chit Funds and Nidhis, which do finance consumption, trade and industry dominate.

Swaminathan (1991) reports that in the two villages she surveys, the share of production loans in the portfolio of lenders is 48.5% and 62.8%. The higher share of production loans is in Gokalipuram, which has the higher interest rates (above 36% for all except the richest group of borrowers).

Ghate (1992) also concludes that the bulk of informal credit goes to finance trade and production.

Murshid (1992) studies Dhaner Upore (cash for kind) loans (you get some amount in rice now and repay some amount in rice later) and argues that most loans in his sample are production loans despite the fact that the interest rate is 40% for a 3-5 month loan period.

Gill and Singh (1997) report that the bulk (63.03%) of borrowing from the informal sector goes to finance production. This proportion is lower for the landless laborers but it is an non-negligible fraction (36%).

6. **Rich people borrow more and pay lower rates of interest; more generally it appears that those who borrow more pay lower interest rates:**

   Ghatak (1976) correlates asset category with borrowing/debt in the *All India Rural Credit Survey* data and finds a strong positive relationship.

   Timberg and Aiyar (1984) report that some of the Shikarpuri and Rastogi lenders set a credit limit that is proportional to the borrower's net worth: Several lenders said that they would lend no more than 25% of the borrower's net worth, though another said he would lend up to 33%.

The "Summary Report on Informal Credit Markets in India" (Dasgupta, 1989) tells us that in their rural sample, landless laborers paid much higher rates (ranging from 28-125%) than cultivators (who paid between 21 and 40%). Moreover, Table 15.9 in that report clearly shows that the average interest rate declines with loan size (from a maximum of 44% to a minimum of 24%). The relation between asset category and interest rate paid is less clear in their data but it remains that the second poorest group (those with assets in the range Rs 5,000-10,000) pays the highest average rate (120%) and the richest (those with more than Rs 100,000) pay the lowest rate (24%).

Swaminathan (1991) finds a strong negative relation between the value of the borrower's land assets and the interest rate he faces: The poorest (those with no land assets) pay 44.9% in one village and 45.4% in the other, while the rich (those with land valued at more than Rs 50,000) pay 16.9% and 24.2% in the corresponding villages.

Ghate (1992) notes that the interest rate on very small loans in Bangladesh tends to be very high (Taka 10 per week on a loan of Taka 500, or 86% per annum).

Gill and Singh (1997) show that the correlation between loan size and the interest rate is negative after controlling for the wealth of the borrower, and that the correlation between the wealth of the borrower and loan size is negative after controlling for loan size. They also find a positive relation between the borrower's wealth and the loan he gets.

3 A simple model of the credit market

- Loan repayment is imperfectly enforceable.

- Suppose k dollars invested yields a gross return $F(k)$ and that the gross interest rate is r. A borrower who has a wealth of w and invests k will need to borrow $k - w$. He is supposed to repay $(k - w)r$ at the end of the period.

- But by expending some resources, which we assume to be proportional to the size of the investment, he can avoid repayment altogether. We denote the constant of proportionality by η and assume that it is less than the cost of capital, ρ.

- Lenders will only provide finance up to the point where the borrower has the incentive to repay: this requires $F(k) - r(k-w) \geq F(k) - \eta k$ which gives us:

$$\frac{k}{w} = \frac{r}{r-\eta} \equiv \lambda(r,\eta).$$

- Firms are credit rationed. They cannot borrow as much as they want.

- The amount you can borrow is increasing in your wealth and your η but decreasing in the interest rate.

- The interest rate is equal to the cost of capital. It obviously does not vary across borrowers.

- This is a handy model but does not fit the facts.

3.1 Extending the model: 1

- It is natural to assume that the lender needs to spend resources in order to make the borrower want to repay. In other words, $\eta = 0$ unless the lender spends some resources.

- First let monitoring cost be linear in the amount borrowed: $\phi(k - w)$.

- In this case

$$\begin{aligned} r(k-w) &= \rho(k-w) + \phi(k-w) \\ r &= \rho + \phi \end{aligned}$$

- r will only vary to the extent that ϕ or ρ varies.

3.2 Extending the model: 2

- Let monitoring be a variable cost, ϕ per unit of $\eta k, i.e.$ the cost does not depend on the amount borrowed but on amount invested.

- Under the assumption of competition, the lender just breaks even:

$$r(k-w) = \rho(k-w) + \phi\eta k$$

- For any credit constrained borrower, $k = \frac{r}{r-\eta}w$, which implies that

$$r = \rho + \phi r = \frac{\rho}{1-\phi}.$$

- Aleem calculates ϕr to be 50 cents per dollar lent on average, easily explaining the gap between the 32.5% cost of capital and the 78.5% average interest rate in this data.

- For this ϕ needs to be about 0.6
- Does not explain exclusion

3.3 Extending the model: 3

- Let the monitoring cost be a fixed cost ϕ

- Then the lender's zero profit condition is

$$r(k-w) = \rho(k-w) + \phi$$

- In the model without default, the borrower's IC constraint is now given by

$$r(k-w) = \eta k$$

which together give us

$$\rho(k-w) + \phi = \eta k$$

- We can rewrite this in the form $k = \frac{\rho w - \phi}{\rho - \eta}$. What if $\rho w < \phi$?Is this necessarily more than w?

- This implies that

$$r = \rho + \frac{\phi(\rho - \eta)}{\eta w - \phi}$$

- Multiplier property.

Implications of the model

- Can explain a large wedge between the cost of capital and the interest rate and by implication a very high monitoring cost.

- The interest rate can be very sensitive to the cost of capital and the monitoring cost, if 1-ϕ is small

- The interest rate will be especially sensitive where the interest rate is high relative to the cost of capital

- However we do not explain equilibrium default.

3.4 Some Policy Implications

- What is the total amount lent?

- In the model without default, the borrower's IC constraint is now given by

$$r(k-w) = \eta k$$

while the lender's zero profit condition is

$$r(k-w) = \rho(k-w) + \phi$$

which together give us

$$\rho(k-w) + \phi = \eta k$$

or

$$k = \frac{\rho w - \phi}{\rho - \eta}$$

- One dollar subsidy to monitoring costs reduces ϕ by ρ dollars (since we assume monitoring costs are paid at the end of the period) which increases the amount of resources going to the poor by $\frac{\rho}{\rho-\eta} > 1$ dollars.

- Keeping the interest rate fixed, the effect of \$1 subsidy would have been $\frac{r}{r-\eta} < \frac{\rho}{\rho-\eta}$. The mutiplier adds to the leverage, especially when monitoring is expensive.

- Cutting monitoring costs is the raison d'etre of the micro-credit movement.

- Note however that one dollar subsidy to wealth (w) would have the same effect.

- However this is only true for those who have $\eta w - \phi > 0$.

- Those who have $\eta w - \phi < 0$, start out unable to borrow.

- For these people a wealth subsidy dominates a monitoring cost subsidy.

- But selection issues favor a subsidy to ϕ (or equivalently to ρ).

- This may be why some micro-credit organizations insist on savings as a way into borrowing (especially under the self-help group model): Helping them save may be way to subsidize building wealth.

3.5 Is asymmetric information a problem? Observing unobservables (Karlan-Zinman)

- It is no longer controversial that credit markets are imperfect.

- The question is to understand the exact technology of lending, since policy implications depend on our understanding of this technology.

- This where "observing unobservables fits in"

- Experimental approach to identifying distortions in the credit market:

- 58000 thousand "good" clients of a South African bank: invited by mail to get a new loan.

3.5.1 The question

- Three interest rate effects:

  Adverse selection

  Repayment burden

  Moral hazard

- A design to separate them (Fig 1, Fig 2):

  Different offer rates

  Different contract rates

  Different length of potential contract

- Size of experimental variation (Fig 3)

3.5.2 Results(Table 3):

- Adverse selection for women

  Moral hazard for men

- Why is the contract rate effect so weak?

  Conservative choice of the original lending amount?

  Absence of moral hazard

- Why is the future interest rate effect stronger?

4 The Banking Channel

4.0.3 The banking channel

- Banks take money from depositors and relend them to firms

- Banks have very low cost of capital (because the average person keeps his savings there), they tend to "uptight" lenders

  Because they have to have enough liquidity all the time

  Because they are heavily regulated

  Because they use agents to do the lending and agency problems may be very serious.

5 A model of the banking channel

- Bankss are supposed to channel the deposits to the right people?

- Do they?

- Why may

5.1 The model

Firms:

- Firms in this economy come in two types–H and L–in proportions p and $1 - p$.

- Their production function is as follows: with probability $p(k, \theta_i)$, where $i = L, H$, $\theta_H > \theta_L$ and $p(k, \theta_i)$ is a concave function of k, the firm produces μk. With probability $1-$ $p(k, \theta_i)$ the firm produces 0.

- Assume that $p_k(k, \theta_i) < 0$, i.e. bigger projects are more likely to fail, that $p(k, \theta_H) > p(k, \theta_L)$ which defines the type H firm to be the more productive firm, and that $0 > p_k(k, \theta_H) > p_k(k, \theta_L)$.

- Finally assume that $kp(k, \theta_i)$ is an increasing but concave function of k.

- Firms live for two periods and produce in both– the shock to its productivity is drawn independently from the same distribution in both periods.

- Firms do not save and have no equity, so in each period, production is entirely financed by credit and the current output is all they have to repay any loans.

- No firm-side moral hazard: firms never deliberately default. However if they do not have enough money they have no choice but to default.

Bank:

- There is a bank that is the sole source of capital to this population of firms.

- To simplify matters we will always assume that banks set a single interest rate r for all loans.

- The cost of bank capital is $\rho < r$.

- Lending decisions are made by bankers who work for the bank but maximize their own expected earnings (rather than the bank's earnings), which in turn are a function of the incentives that the banks set for them.

Bankers:

- Each new firm get allocated to a banker who works with it for the next two periods.

- Bankers get paid an amount C per unit they lend.

- They also get punished for any loans that they made that were ultimately defaulted upon. The punishment is given by $\tau F(k)$ where k is size of the loan, F is an increasing function and τ is a shift parameter which measures the intensity of the monitoring regimes.

- The fact that both C and $F(k)$ are taken as being exogenous reflects an incomplete contracts approach

- Risk-Neutrality on all sides

Information:

- In the first period neither the firm nor the banker knows the firm's type.

- At the end of the first period they both observe the firm's output and update accordingly using the fact they both know how much the firm had invested:

- Successful firms are going to be type H with probability

$$p_s(k) = \frac{pp(k, \theta_H)}{pp(k, \theta_H) + (1 - p)p(k, \theta_L)}$$

  Since $0 > p_k(k, \theta_H) > p_k(k, \theta_L)$ and $p(k, \theta_H) > p(k, \theta_L)$, $\frac{p(k,\theta_H)}{p(k,\theta_L)}$ is increasing in k, which is why p_s is increasing in k.

- Likewise the probability that a failed firm is type H, p_f, is declining in k.

- Informativeness goes up with k.

5.1.1 Lending to a successful old firm

- Suppose that a firm is in the second period of its life, was able to repay its first period loan and is of type H with probability p'.

- Then the amount the banker would want to lend to that firm in the second period is given by maximizing

$$
\begin{aligned}
&Ck_2 - [p'(1 - p(k_2, \theta_H)) \\
&+(1 - p')(1 - p(k_2, \theta_L))] \times \\
&(q + (1 - q)\tau)F(k_2)
\end{aligned}
$$

  where we write k_2 to remind ourselves that this is the second period.

- As long as the second order condition for this maximization holds (which we assume), an increase in p' will increase k_2.

- Also k_2 is declining in $q+(1-q)\tau$, which is the expected value of τ, which we will denote by τ^e. More stringent punishments will discourage lending. Hence $k_2 = k_2(p', \tau^e)$, with $k_{2p} > 0$, and $k_{2\tau^e} < 0$.

5.1.2 Lending to a failed old firm

- Consider now the alternative case where everything is the same about the firm except that it has failed in the first period and hence has no output to repay its loan.

- One thing the banker could do is to let the firm declare bankruptcy and then lend to it based on its p' and the value of τ expected for the next period, i.e. $k_2(p', \tau^e)$.

- In this case the banker gets a punishment $F(k_1)$ for the first period default, where k_1 was the first period loan to the firm.

- The alternative is to bail it out: Give it a loan such that it can repay the first period loan (and thereby avoid defaulting) and as well as the second period loan, if it manages to the successful in period 2.

- This means a loan of size $k_b(k_1)$

$$\begin{aligned} \mu(k_b(k_1) - k_1 r) &= k_b(k_1) r \\ k_b(k_1) &= \frac{\mu r k_1}{\mu - r}. \end{aligned}$$

  This would avoid immediate default but create a risk of a bigger default in the future (since $r > 1, \frac{\mu r}{\mu - r} > 1$).

- Assume that this is bigger than $k_2(p', \tau^e)$ (if this is not true, always bailout).

- The basic trade off

  Bailout now avoids definite punishment

  At the cost of postponed probabilistic punishment for a bigger default

  And you get to lend more

- Whether there will be a bailout depends on

  Whether the probability of punishment is going up or down.

  Shape of punishment function

    The evidence from India seems to be that the punishment itself is clearly quite limited. Even among the most publicized cases one out of 76 went to prosecution

  The probability of punishment is increasing with loan size but slowly

5.1.3 Lending to a new firm

- Goes down relatively more when τ is high because there is more uncertainty about them.

5.2 Direct evidence for agency problems (from Hertzberg, Liberti and Paravasini, 2007)

- Take the model above of the banking channel and assume that a fraction of the firms survive to a third year but are transferred to another lender.

- Also assume that in order to justify lending a lot to an old firm it must be rated highly.

- So the bailed out firms were rated highly at the begining of the 2nd period.

- However the loan officer who takes over observes the history of what happened and can infer the borrower's likely type.

- He has no reason to give a big loan to those bailed out firms that succeeded in the second period and hence survive. He will want to cut the loan that they are getting

- He will surely down-grade them in terms of their rating.

- Anticipating this discrepancy, the first loan officer will start down-grading them from the end of the secodn period.

- And the ratings given at that time will be much better at predicting borrower perforamnce

- This is the prediction they test using Argentine bank data

- In this bank loans are supposed to get transferred every 3 years.

- They find that ratings crash at the end of that period and their predictive power rises.

5.3 More direct evidence (based on Banerjee-Cole-Duflo (2008)

The effect of CVC investigations on Indian banks

- Sharp increase in probability of an investigation followed by decline over 3 years

- Should lead to

  Slowing in lending

  Especially to young firms

  Possibility of bailouts

- Observable implications

  Immediate cut in lending

But the fall in lending persists

Because of bailouts slow down the initial cut

Less information generated about small firms

Default might go up later.

5.4 Are bank clients credit-constrained?(based on Banerjee-Duflo (2002)

- Access to banks is often used as a measure of financial development

- Only relatively privileged firms have access to bank credit.

- However as we have already seen, there are good reasons why bank clients may not get as much credit as they want from the bank

  This does not mean that they are credit constrained: they might get the extra credit they want elsewhere.

5.4.1 An empirical approach to credit constraints

- How do we know whether a firm is credit constrained?

- We need to know its marginal product of capital, but how can we estimate the production function?

- A natural experiment approach

- Indian banks, both private and public, are required to lend 40% of their portfolio to the priority sector.

- In January 1998 firms India with fixed capital between Rs. 6.5 million and Rs. 30 million became eligible for (possibly subsidized) priority sector credit from banks. Firms below Rs.6.5 million were already eligible.

- In early 2000, the limit was lowered back to Rs. 10 million.

- We study the impact of newly becoming eligible/ineligible for subsidized credit on the growth rate of borrowing, sales and profits using firm level data that we collected from a single bank.

5.4.2 Theoretical challenge

- The fact that firm absorbs more subsidized credit does not mean that it is credit constrained.

- To be credit constrained you should be willing to borrow more at the interest rate you pay on the marginal dollar you borrow (not necessarily the subsidized rate, which may be infra-marginal).

- Unconstrained firms will use subsidized credit to pay down their existing debt:

  – they only expand production once they only have subsidized debt.

  – their production(sales) will grow slower than their credit.

- Constrained firms will use subsidized credit to expand sales.

5.4.3 Estimation

- We will mainly estimate

$$\begin{aligned} y_{it} - y_{it-1} &= \alpha_y BIG_i + \beta_y POST_t \\ &\quad + \gamma_y BIG_i * POST_t + \epsilon_{yit}, \end{aligned}$$

  for $y =$ logcredit, logrevenue, logprofits, etc; BIG represents newly eligible firms; the dummy $POST$ represents the post January 1998 period or the post January 2000 period.

- We will also estimate the effect of credit on sales or profits by instrumenting credit by $BIG*POST$

- $BIG * POST$ is uncorrelated with the probability of an enhancement in the loan size.

- Strongly correlated with loan size conditional on there being an enhancement.

- Because it is uncorrelated with the probability of enhancement, we can focus on the firms that got an enhancement

Results

- The OLS effect of growth in credit on growth in revenues is essentially zero. Why might this be?

- What do we learn from using the policy shock? Who would be the compliers in our theory?

- Credit to BIG firms grows faster in the POST period (column 2)

- No change in the interest rate (column 3)

- Firms appear to be credit constrained–sales grows almost as fast as credit (column 5) suggesting that they are not using subsidized credit to pay off market borrowings (substitution).

- Sales grows at about the same rate at firms that have no market borrowing and at firms with some market borrowing (Column 5,6), confirming that there is no substitution.

- Profit has an elasticity of 1.8, implying that an extra rupee of credit increased profits net of interest by almost 1.4 rupees.

Conclusion

- Firms are clearly severely credit constrained.

- There is clearly a large wedge between the rates paid to savers and the marginal product of capital

- Marginal product is very high (possibly over 100%) for the set of compliers.

- This does not directly tell us about whether the marginal product is equalized in all uses.

- However it does suggest that people who have wealth would rather invest it than put in the bank, even if the investment is not the most productive.

5.5 Is the marginal product of capital equalized?(from Banerjee-Munshi (2001))

- If the credit market institutions function poorly, people will prefer to lend to relatively unproductive people who they trust, over a highly productive stranger. As a result social conections, as much as productivity, will determine the allocation of capital.

- Exporters in the knitted garment industry in Tirupur, India, belong either to the local Gounder community or are Outsiders. The Gounders are reputed to be cash rich.

- Our strategy is to compare the investment and production behavior of Gounders and Outsiders, based on our survey data. Essentially all we do is to compare means and growth rates, after controlling for cohort effects and years of experience.

Results

- Gounders who start firms start with almost three times as much fixed capital as the Outsiders

- Gounders own more fixed capital, at all levels of experience, though Outsiders catch up over time.

- Gounders own significantly more fixed capital per unit of production, at all levels of experience.

- Initially Gounders produce more, but the Outsiders have faster growth rate and produce signficantly more after being in business for six years or more.

- Within each community those invest more produce more and grow at least as fast.

Conclusions

- The contrast between within community and across communty patterns rules out any obvious exclusively technological explanation of this evidence, suggesting that community-specific factors must play a role.

- We can plausibly rule out community specific factors other than access to capital:

  - Gounders ought to have better access to inputs other than capital (sub-contracting, politically provided inputs). But then Gounders should have been more productive.

  - The fact that Gounders start big may result from better acess to buyers, but why would they invest more even when they produce less?

- We conclude that Gounders invest more despite being less good at making use of their capital.

14.771 Development Economics: Microeconomic Issues and Policy Models
Fall 2008

Savings

ABHIJIT V. BANERJEE

Why don’t the poor save?

- Lack of savings opportunities?
- Data from vegetable vendors in India.
- Simple production function
  - Purchase fruit in the early morning
  - Sell through day
- Basic working capital needs

Fruit Vendor

Photograph of woman selling fruit removed due to copyright restrictions.

Vendors

Table 1-Business Characteristics of sample population			
Detail	**Percentage of respondents**	**Average amount purchased***	**Profits per day***
1. One trip a day to the market- normal days	89.7%	Rs. 1075.3 (589.2)	Rs.110.5 (54.7)
2. twice or more trips a day(total amount purchased per day)	8 %	Rs.707.5 (422.6)	Rs.95.6 (46.1)
3. once in two days trip to the market (amount purchased per trip)	2.3%	Rs. 1034.8 (515.8)	Rs.97.2 (44.3)
4. good days a week	98.9%	Rs. 1666.3 (834.3)	Rs. 186.6 (83.4)
5. festival days	91.5%	Rs. 2580.7 (1543.7)	Rs. 318.2 (187.3)

The Use of Savings

Table 4- Usage of savings products	
Savings product	**Usage by respondents (in %)**
Cash at home	77.5
Cash lent out	5.7
Cash saved with family/friends	1.5
Chit funds	11.2
MFI/SHG	29.2
Bank account	12.8
Gold	74.6

The puzzles: Vendors have debt

Table 3- Meter loans for financing	
1. % of sample size that takes daily loans	**69.4%**
2. % of sample size that takes daily loans for more than 15 days a month	65.7%
3. average number of days in a month that respondent takes a daily loan for working capital	25.8 days
4. average number of years of taking daily loans	**9.5 years**
5.average daily interest rate	**4.9%**
6. % of total meter loan borrowers who borrow from the same moneylender daily	67.7%
7. Average of maximum that can be borrowed as a daily loan	Rs. 4098.6
8. % of meter loan borrowers who feel there is no other way of doing business and the interest is unavoidable	63.8%

Vendors

- Persistent borrowers
- At very high rates (10% per day)
- Stark implication:
  - One less cup of tea a day.
  - In **30 days** will have doubled **income.**
- Significant foregone income

Vendors Problem not unique

- Payday Loans
  - Skiba Tobacman, 18% for loans lasting two weeks
  - People take many loans before defaulting
    - In essence paying the entire amount on their cycle before defaulting
- Many other apparently myopic behaviors
  - Drug adherence

Intertemporal substitution

- Recall basic Euler equation for someone borrowing at rate R

$$u'(c_t) \geq \delta R u'(c_{t+1})$$

- Basic intuition:
  - People can always borrow less and finance out of their own consumption.

Implications of high interest rate

$$u'(c_t) \geq R\delta u'(c_{t+1})$$

- Discount future heavily (δ low) *or*
- Future marginal utility large relative to today
  - Consumption growth large
    - $u'(c_{t+1})$ low so c_{t+1} high
    - Note: this is stronger than saying that marginal product of capital is high.
      - Some existing studies suggest this as well.
  - Particularly sensible for transitory shocks (e.g. health).
    - But examples span even working capital uses (e.g. crop finance)

Understanding Poverty

- To fit these facts current models must assume

Poor are very myopic

or

Poor cannot cut back consumption

or

Poor are quickly becoming non-poor

Or

Poor do not understand compound interest

Testing these Hypotheses

- Experiment (Karlan-Mullainathan)
  - Buyout the debt
  - Provide literacy

		Financial Literacy	
		No	Yes
Debt Buyout	No	1/4	1/4
	Yes	1/4	1/4

Interventions

- Buyout
  - Give a cash grant enough for individuals to buyout their debt
  - Working capital on a good day (gotten from the baseline survey). As high as 3000Rs.
- Training
  - Half day class where we:
    - Worked out how much they've spent in total on interest rate
    - Benefits of cutting down: illustration
    - Discussed what they could have done with the money
    - Brainstorm on ways to cut down

Sites

- Philippines: Follow up surveys occur
  - 2 weeks
  - 6 weeks
  - 10 weeks
- India: Follow up surveys occur
  - 3 months
  - 6 months
  - 12 months

Summary Statistics, Baseline					
	Control	Training	Debt pay-off	Both	Total
	(1)	(2)	(3)	(4)	(5)
Panel A: India					
Thandal Loan	0.620	0.640	0.664	0.672	0.649
	(0.031)	(0.030)	(0.030)	(0.030)	(0.015)
Thandal Loan amount	2838.40	3006.80	3303.80	3458.00	3151.75
	(226.31)	(256.11)	(248.63)	(259.63)	(124.06)
Moneylender loan	0.844	0.804	0.780	0.780	0.802
	(0.023)	(0.025)	(0.026)	(0.026)	(0.013)
Moneylender Loan amount	21948.13	18349.64	21633.74	26477.54	22102.26
	(2110.67)	(1616.54)	(1773.82)	(4219.66)	(1324.53)
Buying goods on credit	0.388	0.380	0.416	0.418	0.400
	(0.031)	(0.031)	(0.031)	(0.031)	(0.016)
Amount of goods bought on credit	747.938	677.947	773.269	771.683	744.075
	(57.057)	(65.627)	(64.582)	(55.487)	(30.351)
Coping mechanism when hit by a negative income shock					
Saving	0.032	0.040	0.024	0.028	0.031
	(0.011)	(0.012)	(0.010)	(0.010)	(0.005)
Borrowing from moneylenders	0.160	0.180	0.184	0.220	0.186
	(0.023)	(0.024)	(0.025)	(0.026)	(0.012)
Borrowing from someone	0.348	0.372	0.324	0.376	0.355
	(0.030)	(0.031)	(0.030)	(0.031)	(0.015)
Means other than borrowing	0.192	0.140	0.132	0.156	0.155
	(0.025)	(0.022)	(0.021)	(0.023)	(0.011)
Total household expenditures in the past month	5688.72	5399.84	5543.02	5516.55	5536.94
	(389.56)	(171.98)	(169.48)	(173.83)	(122.46)
Total food expenditures in the past month	2807.20	2424.40	2428.40	2535.60	2548.90
	(364.00)	(69.39)	(70.01)	(68.39)	(95.80)
Number of observations	250	250	250	250	1000

	Control	Training	Debt pay-off	Both	Total
	(1)	(2)	(3)	(4)	(5)
Panel B: Philippines					
Moneylender loan	0.984	0.968	0.984	0.952	0.972
	(0.016)	(0.023)	(0.016)	(0.027)	(0.010)
Moneylender Loan amount	3658.730	3975.806	3661.290	3711.111	3751.200
	(267.46)	(323.47)	(300.22)	(339.06)	(153.63)
Buying goods on credit	0.333	0.258	0.371	0.270	0.308
	(0.06)	(0.06)	(0.06)	(0.06)	(0.03)
Amount of goods bought on credit	232.667	30.081	356.484	264.127	221.060
	(130.01)	(19.42)	(159.04)	(192.89)	(70.79)
Coping mechanism when hit by a negative income shock					
Saving	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Borrowing from moneylenders	0.032	0.032	0.016	0.032	0.028
	(0.022)	(0.023)	(0.016)	(0.022)	(0.010)
Total household expenditures in the past month	7037.576	7505.524	6012.747	6951.414	6877.756
	(470.68)	(577.11)	(452.18)	(483.02)	(249.73)
Total food expenditures in the past month	4259.690	4297.629	3488.032	4467.582	4130.117
	(327.83)	(227.90)	(269.89)	(315.61)	(145.41)
Number of observations	63	62	62	63	250

Results - Borrowing

	Follow up 1 (2 weeks after the intervention)		Follow up 2 (6 weeks after the intervention)		Follow up 3 (10 weeks after the intervention)	
Specificaion:	Probit	OLS	Probit	OLS	Probit	OLS
Dependent variable:	Moneylender	Log (loan amount)	Moneylender	Log (loan amount)	Moneylender	Log (loan amount)
	(1)	(2)	(3)	(4)	(5)	(6)
Post x Pay off	-0.332***	-0.275*	-0.302**	-0.368**	-0.201*	-0.340**
	(0.126)	(0.164)	(0.122)	(0.151)	(0.112)	(0.149)
Post x Training	0.042	-0.130	0.009	-0.109	0.044	-0.166
	(0.055)	(0.153)	(0.068)	(0.143)	(0.065)	(0.145)
Observations	500	417	500	412	500	404
R-squared	0.323	0.045	0.314	0.06	0.271	0.057
Dep.var.mean	0.834	8.160	0.824	8.167	0.808	8.158

Results - Borrowing

	Followup 1 (3 months)			
Specification	probit	OLS	probit	OLS
Dependent Variable	Thandal loan	Log(thandal loan amount)	Moneylender loan	Log(Moneylender loan)
	(1)	(3)	(2)	(4)
Post x Training	-0.038	-0.288	0.045	-0.030
	(0.045)	(0.367)	(0.030)	(0.285)
Post x Debt pay off	-0.103**	-0.856**	-0.027	-0.370
	(0.045)	(0.367)	(0.038)	(0.285)
Observations	2000	2000	2000	2000
R-squared	0.013	0.01	0.165	0.19
Dep.Var.Mean	0.591	4.905	0.830	7.506

Results-Borrowing

	Followup 2			
Specification	probit	OLS	probit	OLS
Dependent Variable	Thandal loan	Log(thandal loan amount)	Moneylender loan	Log(Moneylender loan)
	(5)	(7)	(6)	(8)
Post x Training	-0.015	-0.119	0.068*	0.075
	(0.047)	(0.334)	(0.040)	(0.284)
Post x Debt pay off	-0.021	-0.263	-0.015	-0.142
	(0.047)	(0.334)	(0.047)	(0.284)
Observations	2000	2000	2000	2000
R-squared	0.121	0.17	0.281	0.47
Dep.Var.Mean	0.449	3.649	0.729	6.472

Thandal Loans

	Followup 1 only		Followup 2 only	
Dependent Variable	Bought goods on credit	amount bought on credit	Bought goods on credit	amount bought on credit
	(1)	(2)	(3)	(4)
Post	-0.154***	-42.159	-0.205***	-96.898***
	(0.023)	(32.192)	(0.025)	(30.019)
Post x Training	-0.012	11.198	-0.003	12.407
	(0.028)	(34.848)	(0.030)	(33.826)
Post x Debt pay off	-0.078***	-106.116***	-0.034	-65.613*
	(0.028)	(34.874)	(0.030)	(33.839)
Observations	1940	2000	1922	2000
R-squared	0.185	0.057	0.200	0.070
Dep.Var Mean	0.301	244.86	0.295	229.598

How are people slipping?

- What drives the long term fall?
- In India we see the biggest fall
- There is some *very preliminary* evidence
  - Question: How did you cope with shocks last month?

What does this tell us

- Cannot be physical inability to save
- Cannot be that much impatience
  - At 10% per day, 1 dollar today is worth less than 1/50 of cent in 3 months
  - Also they buy durables, marry their daughters
  - It could all be borrowing but why do they repay? After all the future credit is worth nothing to them
  - How do they manage to remain in a ROSCA year after year?.
-

What does this tell us

- Probably not a lack of understanding
- Particular kind of self-control problem?
- Can we learn something from how they fall back?

Results – Coping With Shocks by..

	Followup 1 only			
Dependent Variable	Savings	Loan	Any Loan	Savings or Non-Loan Source
	(1)	(2)	(3)	(4)
Post x Training	-0.027	-0.033	-0.055	0.002
	(0.020)	(0.035)	(0.042)	(0.036)
Post x Debt pay off	0.074**	-0.081**	-0.060	0.083**
	(0.034)	(0.033)	(0.042)	(0.040)
Observations	2000	2000	2000	2000
R-squared	0.078	0.010	0.005	0.015
Dep.Var.Mean	0.081	0.220	0.375	0.195

Results- Coping with Shocks by…

	Followup 2 only			
Specification Dependent Variable	Savings	Loan	Any Loan	Savings or Non-Loan Source
	(5)	(6)	(7)	(8)
Post x Training	-0.016	-0.058*	-0.050	0.005
	(0.018)	(0.034)	(0.042)	(0.032)
Post x Debt pay off	0.019	-0.035	0.011	0.043
	(0.024)	(0.036)	(0.044)	(0.035)
Observations	2000	2000	2000	2000
R-squared	0.035	0.011	0.003	0.002
Dep.Var.Mean	0.058	0.226	0.381	0.150

Modeling myopia

- Two periods in most examples
- Two types of index goods: *x* and *z*
  - x consumption: no time inconsistency
  - z consumption: only present selves like it
- Instantaneous utility in each period $u(x) + v(z)$
- Period 1's decision utility:

$$u(x^1) + v(z^1) + \delta u(x^2)$$

- Income each period y^t and initial wealth w^0
- Production function $f()$. Sometimes for simplicity will just assume rate of return R

Generalized Euler Equation

- Traditional Euler Equation:

$$u'(c_t) = \delta f'(w_t) u'(c_{t+1})$$

- Generalized Euler Equation

$$u'(c_t) = \delta f'(w_t) u'(c_{t+1})[1 - z'(c_{t+1})]$$

- Temptation tax:
  - Every dollar transferred into the future is “taxed” by temptations; future selves will waste some of it.

Poverty and Myopia

- Two forms of "myopia": δ and z'(w)
- Original puzzle
  - Third explanation: myopia in the form of high z'(w).
- Why is this different?
  - Because z'(w) can vary systematically with w
  - Individuals can control *the value of z'(w)* they face and hence the tax.
  - All our results come from this.

The shape of temptation

- Two important cases:
  - $z'(c)$ constant (Non Declining temptation)
    - Rich and poor face similar time inconsistency problems
    - Includes case of $z'(c) = 0$
  - $z'(c)$ declining
    - Rich face less time inconsistency problems

What does this framework give us?

- Demand for commitment: not just by some “cold” self: Size effect
  - Ashraf, Karlan and Zin (“Tying Odysseus to the Mast”)
  - ROSCA participation
    - Anderson and Baland think its spouse control
  - Microfinance participation
  - Excess purchase of durables
- Aspiration effect: when the future looks better people might save more
- Lack of buffer stocks against income risk
  - Rosenzweig-Wolpin

Rosenzweig-Wolpin

- Bullocks: draught animal in India: Usually a pair of them used for tilling land
- They jointly estimate a linear production function: f
  - Farm profits = A. #bullocks + B. pump + C. #bullocks.pump + village-year dummy+ e
- And a Stone-Geary utility function
- Assume that the shock is realized before farm inputs are put in: separability
- Using The ICRISAT panel. 30 farmers, 9 years

Conclude

- That bullocks are very profitable—cost 1000 rupees. Yield 1400 rupees more profits (but cost of feeding)
- So are pumps
- Yet 31% have ever owned a pump
- And 10% sold a bullock last year. More sales in bad weather years
- Durables are being used for consumption smoothing.

Implications of constant z'

- Useful applied insights
  - No different than applying standard models (e.g. hyperbolic)

Example applications

- Demand for Commitment
  - SEED, ROSCAs
- Purchase of Durables
  - Suppose durables provide fixed x utility
    - Individuals willingness to pay for durables will be

$$p = \frac{u_d}{u'(c_t)}(1+\delta)$$

    - If discount factors on consumption or investment data assuming a traditional Euler equation, individuals will appear to over-demand durables relative to investments

$$p = \frac{u_d}{u'(c_t)}\left(1+\frac{\hat{\delta}}{(1-z'(c_t))}\right)$$

Demand for durables

- By over-investing enough in durables the current decision-maker locks in future x consumption (assuming that durables generate u consumption.

What is a Temptation?

- Demand for commitment devices also tells us potentially what is a x-good?
  - People would only save up (in a commitment device or otherwise) to buy an x-good.

Clients' Specific Savings Goals

	Frequency	Percent
Christmas/Birthday/Celebration/Graduation	97	48.0%
Education	42	20.8%
House/Lot construction and purchase	21	10.4%
Capital for business	20	9.9%
Purchase or maintenance of Machine/Automobile/Appliance	8	4.0%
Agricultural Financing/Investing/Maintenance	4	2.0%
Vacation/Travel	4	2.0%
Personal Needs/Future Expenses	3	1.5%
Did not report reason for saving	2	1.0%
Medical	1	0.5%
Total	202	100.0%
Data-based goals	140	69.3%
Amount-based goals	62	30.7%
Total	202	100.0%
Bought ganansiya box	167	82.7%
Did not buy ganansiya box	35	17.3%
Total	202	100.0%

Figure by MIT OpenCourseWare.

Declining Temptation

- Really where model can be more insightful

- Why might temptations decline?
  - Basic temptations—sugar, fat, addictions—dealt with first
  - Supply: aimed at average income
- Ultimately an empirical question
  - Here, we draw out the consequences.
  - Will talk about direct tests of z’ as well

- Why not consider z’ increasing?
  - Uninteresting: strong convergence

Demand for Commitment

- This implies that individuals will demand specific types of commitment accounts
  - SEED size-based goals (Ashraf, Karlan and Yin)
  - To explain time-based would need to assume that $u'(x)$ is particularly high relative $v'(z)$ at certain periods.
  - Size element of ROSCAs

Outline

- Attributions of impatience
- Impact of future income
- Poverty trap
- Response to uncertainty
- Investment features
- Role of credit
- Money Lender
- Testing this model

Outline

- **Attributions of impatience**
- Impact of future income
- Poverty trap
- Response to uncertainty
- Investment features
- Role of credit
- Money Lender
- Testing this model

Attributions of Impatience

- Suppose we observe a population of individuals with a distribution of δ and initial wealth which have correlation ρ. All have the same u(x) and v(z).
- Suppose an econometrician estimates on this data a time consistent utility function for total consumption and a distribution of δ
- Estimated discount factor of individual *i*

$$\hat{\delta}_i = \delta_i(1 - z'(c_i))$$

Attributions of Impatience

- The poor will look more impatient

$$\mathrm{cov}(\hat{\delta}_i, w_i) > \mathrm{cov}(\delta_i, w_i)$$

- Intuition: Poor face higher z'(c)
  - Those with higher z'(c) tend to consume more today.
  - As a result the econometrician, who assumes exponential discounting, will attribute that steeper consumption profile to a smaller discount factor.
  - But since this effect is bigger for the poor than the rich, the misattribution of greater impatience will be larger for the poor and will induce a positive correlation discount factors and income, even if none existed.
- The poor face bigger temptations

Outline

- Attributions of impatience
- **Impact of future income**
- Poverty trap
- Response to uncertainty
- Investment features
- Role of credit
- Money Lender
- Testing this model

Future income

- *Proposition* Assume that second period income, y_2; is deterministic. If temptations are not declining, period 1 consumption increases with period 2 income

$$\frac{dc_1}{dy_2} > 0$$

If temptations are declining then there exist utility functions for which there is a range of y_2, where consumption in period 1 decreases with income in period 2

$$\frac{dc_1}{dy_2} < 0$$

Moreover we will only observe this pattern for people for whom y_1 and y_2 are sufficiently small.

Intuition

- Consider the Euler equation

$$u'(c_2) = \delta f'(y_1 - c_1)u'(c_2)[1 - z'(c_2)]$$

- If consumption today doesn't change with y_2 then right hand side:
  - Goes down because $u'(c_2)$ rises.
  - Could go up if $z'(c_2)$ falls
- With constant temptation first effect implies c_1 must rise.
- With non-constant temptation, there are two effects.

Intuition

- Aspiration effect
  - If the future looks bleak, there is little point in saving.
- This is the core of most of our propositions below

Future Income

- Another intuition: Suppose an individual has a time consistent utility function

$$u(c_1) + \delta u(c_2)$$

But has a strange investment technology

$$\tilde{f}(\bullet) = f(\bullet)[1 - z'(w_2 + y_2)]$$

- Thus an increase in y_2 has two effects:
  - Consumption smoothing as before
  - An increase in the investment efficiency
- This intuition will help us think about several of the examples below.

Outline

- Attributions of impatience
- Impact of future income
- **Poverty trap**
- Response to uncertainty
- Investment features
- Role of credit
- Money Lender
- Testing this model

Poverty Traps

- *Proposition* Suppose there is no uncertainty. Then when temptations are not declining, c_2 is continuous in initial income y_1. When temptations are declining, however, a poverty trap can emerge: for some parameters, there will exist K such that c_2 jumps discontinuously at K. Moreover, $u(x_1) + \delta u(x_2)$ and $u(x_1) + v(z_1) + \delta [u(x_2) + v(z_2)]$ are both discontinuous in y_1.
- Notice: no increasing returns (or even credit constraints)

Intuition

- Logical consequence of income effect from above.
  - Greater wealth → more to save
  - More to leave behind → Lower *z'(c)*
  - Lower *z'(c)* → Greater incentive to save
- Another intuition:
  - Investment "technology" becomes more efficient

Interpretation

- Poor are penalized by having more of their money “wasted”
  - Dulls their incentive to save
- Multiple periods exaggerates this trap
  - Better behavior by 3 generates better behavior by 2 which generates better behavior 3
  - Generates a strategic incentive to save:
    - Increase *z’()* for future selves and they will strategically save to further increase *z’().*
- Adds nuance to accumulation for lumpy investment
  - At low levels of wealth, accumulation is “leaky” due to temptations

Outline

- Attributions of impatience
- Impact of future income
- Poverty trap
- **Response to uncertainty**
- Investment features
- Role of credit
- Money Lender
- Testing this model

Response to Uncertainty

- Consider now the case where y_2 can be uncertain. We will consider wat happens when uncertainty increases, i.e. the effect of mean preserving spreads of y_2 on c_1.

- Define the indirect utility function

$$w(c) = \max_x u(x) + v(c - x)$$

Response to Uncertainty

- *Proposition* If *w(c)* exhibits prudence and temptations are non-declining, then $c_1$decreases with uncertainty in y_2.
  If *w(c)* exhibits prudence and temptations are declining, then there exist situations where c_1 *increases* with uncertainty in y_2.

Intuition

- Back to asset intuition:
  - Uncertainty in y_2 means that investment return has risk: *z'()* could be low or high. But notice that this risk is badly correlated: pays off most when needed least (high income state)]
- So increased risk:
  - Prudence
  - Higher correlation of investment returns; more risky asset
- Two offsetting effects

Insufficient Buffer Stock Savings

- Very important practical issue:
  - Poor often living on edge
  - Very little buffer stock savings
- Observations
  - In two periods could be practically constrained by range where z'() is actually increasing (starvation)
  - In multiple periods effect is magnified
  - A hidden effect: for those who are near poverty trap threshold, uncertainty can be very good

Example: Payday Loans

- US poor often borrow at very high rates for payday loans
- Note that the problem may not be taking out the loan
  - Faced with shock that could have large consequences, taking loan may be sensible
  - Key problem is lack of saving in the past that brought them to the point where they need a payday loan

Outline

- Attributions of impatience
- Impact of future income
- Poverty trap
- Response to uncertainty
- **Investment features**
- Role of credit
- Money Lender
- Testing this model

Investments

- What does this model imply about the types of investments people will undertake?
- To answer this we consider the following thought experiment.
  - Define a linear investment technology to be defined by $H = (R,s,S)$, where this technology allows an individual to invest any amount between s and S at a linear return R.
  - We will consider someone who has access to H on top of the standard technology
  - Suppose he undertakes some investment in H.
  - Suppose an identical person has access to $H' = (R',s',S')$ and the standard technology
  - What conditions determine whether he will undertake some investment in H' ?

Investments

Proposition If temptations are not declining, then investing in H implies investing in H' as long as $R' \geq R$ and $s \geq s'$. In other words minimum scale and returns summarize the investment decision.

If temptations are declining, then there exist situations where this is not true if $S \geq S'$. In this case, *maximum scale* also determines investment

High Return Investments

- Aspiration effects
  - Unless an investment has a big (in level) change, it doesn't matter.
- Effectively creates minimum scale even in linear investments
- Potentially helps explain high return investments which are *divisible* but are not undertaken
  - Fertilizer (Duflo, Johnson, Kremer)
  - Stocking (Lee, Kremer and Robinson)

Outline

- Attributions of impatience
- Impact of future income
- Poverty trap
- Response to uncertainty
- Investment features
- **Role of credit**
- Money Lender
- Testing this model

Credit

- In all self-control models credit can potentially be very bad
  - Can exaggerate self-control problem
- To understand this, we introduce artifice of *0* period self
  - Does not consume
  - Maximizes $u(x_1) + \delta u(x_2)$
- He chooses whether or not to allow a particular credit option.

Credit

- We will consider the following thought experiment.
  - Define a credit technology to be $C = (R,s,S)$, where an individual can borrow any amount between s and S at a linear cost R.
  - Zero period self has the choice of whether or not to add access to C for period 1 on top of the existing technology
  - Suppose zero allows C.
  - Consider an identical person where 0 must decide whether to allow access to $C' = (R',s',S')$ on top of the existing technology
  - What conditions determine whether zero will allow C'?

Credit

Proposition If temptations are not declining, then allowing C implies allowing C' as long as $R' = R$ and $S \geq S'$ and $s \geq s'$. In other words he might want to place a cap on the maximum loan available.

If temptations are declining, then there exist situations where this is not true. This occurs when $s < s'$. In other words, zero period self will want to place a *floor* on the minimum loan available

Intuition

- Constant temptations ➔ fear is overborrowing
  - Don’t want 1 to take too much.
- Declining temptations ➔ At higher levels, may be more willing to invest.
  - Hence bigger loan may be good
  - And may even want to impose floors
    - Small amounts wasted. When that option is not there, big amount can be invested.
  - Note: Could get same effect if there is constant temptation and lumpy investments.

Implications

- Credit cards
- Micro-finance loans
- Can have different implications for self-control and temptations.

Outline

- Attributions of impatience
- Impact of future income
- Poverty trap
- Response to uncertainty
- Investment features
- Role of credit
- **Money Lender**
- Testing this model

Money Lenders

- Old argument (Bhaduri) on how money lenders can trap individuals in poverty
  - Prevent them from adopting high return investments
  - Why? If the individual gets wealthier he may rely on money lender less
- Problems
  - Coasian: simply charge higher rate for the investment
  - Conceptual: Why would the person borrow less if wealthier

Money Lenders

- Investment decision
  - An amount to be invested in 0.
    - Zero period self only invests, no consumption. Maximizes $u(x_1) + \delta u(x_2)$
  - A second unobservable investment in period 1.
  - Payoff R in period 2 if both investments made.
- Money lender sets interest rate
  - Two rates: R_0 and R_1.
- Define R'_1 to be the rate charged by the money lender when this investment is not available.
- Suppose that at $R_0 = R_1 = R'_1$ both periods would invest.

Money Lenders

Proposition When temptations are non-declining, both periods would continue to invest though the money lender will charge rates above R'_1

If temptations are declining, however, then there exist parameter values where the investment does not take place.

Note: this occurs even though the investment can be made more attractive because of declining temptation.

Money lender problem

- Money lender faces trade-off
  - Financing investment raises total pie
  - Financing investment can increase wealth and thereby decrease desire to borrow
- Increasing interest rates to offset the second effect (the Coasian solution) will
  - Make period 2 self poorer
  - And hence may make period 1 self less likely to invest.
- Gains from trade not fully exploited because period 1 not fully able to commit

Implications

- Related to literature on debt traps
- Creates interesting income dynamics in economies with monopolistic credit
  - Vast majority of money lenders

Outline

- Attributions of impatience
- Impact of future income
- Poverty trap
- Response to uncertainty
- Investment features
- Role of credit
- Money lender
- **Testing this model**

Testing the assumption

- Multiple goods indexed by *i*
  - Each provides x_i and z_i units of non-temptation and temptation goods.
- Make an offer of 1 unit of good *i* today vs. *k* units tomorrow
  - Note would need non-fungibility to do this exercise
    - Always the case ($10 today vs. $15 tomorrow when you have $100 in your pocket).
- Allows us to estimate good specific discount factor: $\hat{d}_i$

Implcations

- Low discount factor goods should have steeper Engel curves
  - Put differently: dollar-weighted average discount factors rise with income

Testing Impatience

- Estimate discount factors as above for money as well as goods known to be high *x* good *j*.
- We predict that

$$\frac{\hat{d}_m}{\hat{d}_i} < 1 = \frac{\hat{d}_m}{\bar{\hat{d}}_i}$$

and that this ratio increases with income.

Psychologically Richer Alternatives

- Same behaviors; different interpretations
  - Rich are fallible; poor are equally fallible
  - Attention is just greater on fallibility of rich
- Different challenges; same basic psychology
  - Will work through one model carefully
- **Different challenges; different psychology**
  - Mullainathan-Shafir

Example from mental accounting

- Imagine that a friend goes to buy an appliance priced at $100($500/$1000). Although the store's prices are good, the clerk informs your friend that a store 45 minutes away offers the same item on sale for $50 less. Would you advise your friend to travel to the other store to save $50 on the $100($500/$1000) expense?

(Crystal Hall)

Percent traveling to save $50

	$100	$500	$1000
HI (*N* = 76)	54	39	17
LI (*N* = 47)	76	73	87

14.771 Development Economics: Microeconomic Issues and Policy Models
Fall 2008

1 Agriculture: The efficiency of land use

- Share of agriculture in employment is close to 50% for the world as a whole (50% in China, 57% in India).

- Is land used efficiently?

1.0.1 Farm size and productivity: observed relationship

- Farm size productivity differences: see table.

- Profit-Wealth ratios and weather variability (monsoon outset is a measure of the risk faced by the farmer): see figure

- The Profit-Wealth ratio is always greater for small farmers

- Small farmers' profits are hurt much more by uncertainty than large farmers'

1.0.2 Why is this surprising?

- Arguments for increasing returns (the opposite relationship)
  - Technology with fixed costs (tractors, etc..)
  - Larger farmers have better access to capital
  - Larger farmers have better access to politically allocated inputs (evidence from Africa in a book by Bates "Market and states in tropical Africa").
  - The best farmer will have more land...
- Mitigating factors:
  - Rental markets in farm machinery
  - Technological change in not very rapid. Saviness not that important.

1.0.3 What could be going on: Arguments for decreasing returns

- – Agency problems: large farms are cultivated by hired labor, which has fewer incentive to work hard. Small farms are owner cultivated. $\Rightarrow$ Redistributing land will create more owner cultivated land which will be more productive.
  – But why cannot the owner of the land not give the right incentive to the farmers?

1.0.4 Different potential explanations for the observed inverse productivity relationship:

- Differences in land quality

- Differences in farmer characteristics

- Incentive Problems

Problem with the observed relationship: all of this could be going on... How can we separate these different effects.

1.0.5 Evidence: Study by Biswanger and Rosenzweig

- Using ICRISAT data: very detailed panel (repeated observation for every household) data from India.

- Some individuals cultivate both an owner-operated plot and a rented plot.

- Biswanger and Rosenzweig compare the inputs they apply on their own plot and the rented plots, and the overall productivity of both plots.

$$\Pi_{ij} = \alpha + \beta R_{ij} + \eta_i + \upsilon_{ij},$$

- where Π_{ij} is farmer's i outcome (profit, investment) on plot j, and R_{ij} indicate whether the plot is rented. η_i is the unoberved (but fixed) characteristics of the farmers (risk aversion, quality, etc...). We think that η_i and R_{ij} may be correlated, but, for a minute, not υ_{ij} and R_{ij}. What can we do?

- Control for the individual fixed effect to compare plots within individual's. So for example, for all the farmers that cultivate two plots of land, we can run the regression:

$$\Pi_{i2} - \Pi_{i1} = \beta(R_{i2} - R_{i1}) + \upsilon_{i2} - \upsilon_{i1},$$

- The individual fixed effect is gone!

Biswanger and Rosenzweig find a strong negative β. What does this suggest? What could be the remaining problem?

1.0.6 More evidence: Shaban (1987)

- Uses the same data, but controls in addition for plot quality.

- He finds that individual work 40% more on their own land (controlling for land size) and that the productivity is 15% to 30% higher on own land than on rented land (with or without controling for land quality).

- On balance, the evidence suggests that the inefficiency comes from incentive problems.

1.1 Incentive problems: A simple model of sharecropping

- There is a landlord who own two plots of land but can only cultivate 1
- He hires a tenant to farm the other plot
- Cultivation effort is denoted by e.
- The landlord cannot observe e.
- Effort is costly to the tenant: $\frac{1}{2}ce^2$
- Two things can happen:
  - with probability e: Output is H

- with probability e: Output is 0

- The tenant has outside option $\underline{w}$

- The tenant and landlord write a contract which specifies a payment to the tenant
  - a payment h if output is H
  - a payment l if the output is 0

1.1.1 What does the landlord choose for e?

- Maximize $eH + (1 - e)0 - \frac{1}{2}ce^2$
  - What is the solution?
  - Why?

1.1.2 No Limited Liability

- Work sequentially: given h and l, what is the tenant's effort? Tenants want to maximize income minus the cost of effort: $eh + (1 - e)l - \frac{1}{2}ce^2$

- What is the solution for e given h and l?

- How do we need to fix h and l to incite the tenant to choose the optimal effort $\frac{H}{c}$?

- $l=$

- $h=$

- This contract is a *fixed rent contract*.

- How is the rent, R, chosen?

- Tenant has to agree to work with landlord: he has to receive at least $\underline{w}$. $\rightarrow$ exercise: calculate R

1.1.3 Limited Liability

- Imagine that the tenant cannot receive negative payment: *limited liability*.

- What will l be?

- What will e be?

- What will the output be?

- How does it move with h?

- Maximization problem of the landlord: Maximize his income.

$$\max \ e_{tenant}[H - h]$$

$$\max \ \frac{h}{c}[H - h]$$

  - What is the optimal h now?
  - What is the output?
  - How does the output compare to the optimal output?
  - What is the difference h - l?
  - How does it compare to the case without limited liability?
  - Why is the effort smaller than the optimal effort?

1.1.4 Outside Option

Remember that the tenant can choose to work somewhere else and will receive a utility $\underline{w}$. How does it modify the contract chosen above?

- Tenant's utility under the contract:

$$\frac{h}{2}h - \frac{1}{2}c\left(\frac{h}{c}\right)^2 = \frac{1}{2}\frac{h^2}{c} = \frac{1}{8}\frac{H^2}{c}$$

if $\frac{1}{8}\frac{H^2}{c} \geq \underline{w}$, they can choose this contract: Is there anuthing strange about this contract?

if $\frac{1}{8}\frac{H^2}{c} < \underline{w}$, they have to pick a contract which will give at least $\underline{w}$ to the tenant Pick h such that:

$$\frac{1}{2}\frac{h^2}{c} = \underline{w}$$

$h=$

$e=$

- e is always an increasing function of $\underline{w}$

- output is always an increasing function of $\underline{w}$

$\longrightarrow$ increasing the tenant outside option increases productivity

1.1.5 Implications

What does that imply for

- The effort chosen by the tenant versus the landlord?

  The relation between land-size and productivity assuming that some people own 1 unit of and some 2?

  The effect of redistributing land on productivity?

1.2 Risk, insurance and tenancy

- We continue in the world of the previous lecture:

  Success is 1, failure is 0

  The probability of success is e.

  However we now add risk-aversion

- First assume that output is publicly verifiable.

- In that case insurance is provided by the "world market" which is assumed to be risk neutral

  – Landlord's do not need to insure their tenants or vice versa.

  – Therefore we can assume that the tenancy contract is always a fixed rent contract at rent R.

– Then the landlord and the tenant gets insurance on the market

- Suppose there are people who own 0, 1 or 2 pieces of land.

- However people can cultivate only one.

- Then someone who has 2 plots will need to get at least one tenant.

- If he gets one tenant, his earnings are $R + \omega + 1$ when its a success and $R + \omega$ when it is failure, where R is the rent and ω is his non-labor income.

- Then he gets insurance:

The insurer is also concerned about moral hazard since she does not observe effort. Suppose the insurance contract gives him y_H when he succeeds and y_L when he fails

The landlord chooses effort based on the implied incentives

- The insurance contract maximizes

$$\begin{aligned} & eV(y_H, e) + (1-e)V(y_L, e) \\ & st\ e \text{ maximizes } eV(y_H, e) + (1-e)V(y_L, e) \\ R + \omega + e \ = \ & ey_H + (1-e)y_L \end{aligned}$$

- For the tenant it is the same problem, except that he has to pay R and has an outside income ω'. His insurance contract maximizes

$$\begin{aligned} & eV(y_H, e) + (1-e)V(y_L, e) \\ & st\ e \text{ maximizes } eV(y_H, e) + (1-e)V(y_L, e) \\ -R + \omega' + e \ = \ & ey_H + (1-e)y_L \end{aligned}$$

1.2.1 A convenient example: $V(\cdot) = -\exp[-\lambda(y - \frac{1}{2}ce^2)]$

- Writing $r = y_H - y_L$ we can rewrite the maximand

$$\max - \exp[-\lambda(y_L)]\{-e \exp[-\lambda(r - \frac{1}{2}ce^2)] \\ -(1-e)\exp[-\lambda(-\frac{1}{2}ce^2)\}$$

- Notice that the choice of e does not depend on y_L.

- You can raise e by raising r.

- The marginal cost of raising r also does not depend on y_L.

- Therefore we can separate the instruments: use r to give incentives and use y_L to set the levels.

What does that imply for

- The effort chosen by the tenant versus the landlord?

  Will the landlord get a tenant on the last plot?

    what is the role of the outside option?

  The relation between land-size and productivity?

  The effect of redistributing land?

1.2.2 Does that mean that with diminishing absolute risk aversion landlords who have one plot will necessarily be more productive on average than landlord's who have two plots as long as tenants poorer than either one as long as those with one or more plot do not have too good an outside option?

- Not necessarily

- The richer landlord might take more risk on plot 1 even if the tenant takes less risk on plot 2

- More subtle effect: the income effect on the willingness to take risks makes tenants less willing to bear risk, but the insurance company also knows that a smaller amount of risk is enough to motivate a tenant. The net effect on effort is ambiguous.

An example

- Assume two choices for $e = 0, e'$.

- The cost of e' : $\frac{1}{2}ce'^2 = E$

- Optimal insurance contract (y_H, y_L) such that

$$e'V(y_H) + (1 - e')V(y_L) - E = V(y_L)$$

1. which implies $e'[V(y_H) - V(y_L)] = E$
2. and $\omega' + e' = e'y_H + (1 - e')y_L$

- $V'(y_H)\frac{dy_H}{d\omega'} = V'(y_L)\frac{dy_L}{d\omega'}$ from 1.
- $e'\frac{dy_H}{d\omega'} + (1 - e')\frac{dy_L}{d\omega'} = 1$ from 2.

- Solving these two we get

$$
\begin{aligned}
\frac{dy_H}{d\omega'} &= \frac{V'(y_L)}{e'V'(y_L) + (1-e')V'(y_H)} \\
\frac{dy_L}{d\omega'} &= \frac{V'(y_H)}{e'V'(y_L) + (1-e')V'(y_H)}
\end{aligned}
$$

- The utility of a farmer

$$V_F(\omega') = e'V(y_H) + (1-e')V(y_L) - E$$

- From which $\frac{dV_F(\omega')}{d\omega'} = \frac{V'(y_L)V'(y_H)}{e'V'(y_H)+(1-e')V'(y_L)} wp$

- For a pure rentier: $V_R(\omega') = \omega' + R$ (bad notation)

- $\frac{dV_R(\omega')}{d\omega'} = V'(\omega + R)$

- Which of these slopes is greater when $V_F(\omega') = V_R(\omega')$ will determine which curve cuts the other from below and therefore who farms.

- Since $\omega' + R < \overline{y} = e'y_H + (1 - e')y_L$ (risk aversion+cost of effort), $\frac{1}{V'(\omega+2R)} < \frac{1}{V'(\overline{y})}$

- Let $1/V'$ be a convex function. When is this true?

- Then $\frac{e'}{V'(y_H)} + \frac{1-e'}{V'(y_L)} \geq \frac{1}{V'(\overline{y})}$

$> \frac{1}{V'(\omega'+R)}$.

- Which implies $\frac{dV_R(\omega')}{d\omega'} > \frac{dV_F(\omega')}{d\omega'}$

- In other words, the landless may be more willing to farm the plot owners, which will bid up the rent till all land is with tenants, even though the plot owners do not have better outside options.

1.3 Costly verification of output

- Suppose output is costly to verify. Specifically one unit of labor can verify the output of N plots of land.

- Assume everyone has one unit of labor and that this unit is indivisible.

- This unit can either cultivate 1 unit of land or verify output on N plots of land.

- Finally assume that the total output of the N plots that the verifier verifies is unobservable to others and as a result, no one insures the person who verifies output.

- Therefore there are three possible occupations

Tenant: someone who farms 1 unit of land whose output is verified by the landlord and therefore can get insurance

Landlord: someone who verifies output and insures others

Yeoman farmer: someone who gets no insurance but farms 1 unit of land.

- Suppose first risk-aversion is fixed: Everyone is CARA with potentially differing coefficients.

- Who, in terms of risk-aversion will become

  a tenant

  a yeoman farmer

  a landlord

think of the case where one group is quite risk averse and the other is risk-neutral

and there are too few risk-neutral people to cultivate all the land.

- What are the implications for

  Farm size and productivity?

  And the effect of land redistribution?

- Now let risk-aversion vary with wealth and let people have identical utility functions but differing wealth

  who becomes a tenant?

  Farm size productivity?

  The effect of land redistribution?

1.4 Two-sided moral hazard

- A fourth model: two sided moral hazard.

  who becomes a tenant

  Farm size and productivity

  The effect of land redistribution

2 The case for redistributing land

Why redistribute land rather than money? As economists, we tend to think that money is better, since with money, the poor could buy land if they wanted to. So why land reform?

- The giving end: Getting land from the rural rich.

- Common argument (1): land cannot flee to Switzerland, and cannot be hidden: easy to seize

  Yet: Land titles are very sketchy. Formal titles can be quite different from effective control, especially if people have an incentive to do so. Land may not be so easy to take away after all.

- Common argument (2): redistributing land does not create distorsions, since it is a fixed asset (income taxation would reduce labor supply, but land does not).

  Yet: Redistributing land is difficult: it is opposed by landowners who often control important political resources. There are very few instances of large scale land redistribution that did not take place in the midst of massive social upheaval. Land reform may be politically very costly.

- Perhaps we want to tax the rural rich, and not the urban elite (entrepreneurs, etc...), for example because we want to foster industrialization. This does not seem to be important now, since recent examples have favored *market assisted* land reforms, whereby landlords are is compensated out of general tax revenues.

- We may feel that landlords hold onto land because it gives them prestige/political power.

• The receiving end: giving land to the rural poor.

  - Makes them more likely to migrate to the cities. But are cities really too large?

  - Land is an asset: can be used as collateral and therefore may generate extra value (Hernando de Soto). But why would illiquidity help.

– Intrahousehold allocation issues. Perhaps money would be spent by the household head in alcohol etc... whether land will remain in the household. We should make it hard to sell the land then! This may be the most compelling argument in favor of land reform. •

2.1 Does land reform work? Besley-Burgess (2000)

- Few studies of the efficiency consequences of large-scale reforms of property rights: Most reforms have been accompanied by major upheaval and social unrest $\rightarrow$ difficult to separate the effects of the two.

- Besley-Burgess (QJE, 2000): Uses the fact that India has carried out a large number of land reforms under conditions of reasonable social peace.

- The number of reforms is large because reform is a state subject and there are many states. Moreover while every state has some reform, some have had more reforms than others.

- Besley-Burgess create an index of the number of land reforms (actually they classify reforms into four categories: tenancy reforms, abolition of intermediaries, land ceilings, land consolidation). They then count the number of reforms of each type in or before a given year. This gives them an index of "tenancy reforms" for that state for that year, say. They add up all the four indices to get a single land refor m index for that state for that year.

- They then run a regression of the form

$$x_{st} = \alpha_s + \beta_t + \gamma y_{st} + \xi I_{st-4} + \varepsilon_{st}$$

  where x_{st}, is the poverty rate in state s at time t, y_{st} is a set of time and state varying controls and I_{st-4} is the index of the number of reforms till up to 4 years before the current date.

- The results show [SEE TABLE III]

  - land reforms are associated with a decline in poverty.

  - There is no effect on urban poverty.

  - Only tenancy reforms and the abolition of intermediaries had an effect on poverty.

- The results survive when you control for: population growth, agricultural yield lagged four years, health and education expenditures per capita, state taxes/SDP, proressivity of taxes.

- Both progressivity of taxes and State taxes/SDP reduce poverty but the land tenure effect remains intact.

- Concern: Endogeneity of land reforms: Instrument for land refoms using the vote share of "hard left" parties as against the centrist parties. Is this a good instrument?

- Concern: Serial correlation of errors.

2.2 Does land reform work? Banerjee-Gertler-Ghatak (2002)

- Banerjee, Gertler, Ghatak study a *tenancy reform* = improvement in the rights of tenants. It differs from a traditional land reform (redistribution of land). Land is not redistributed. The tenant is offered the *security of tenure* = if he registers, he cannot be evicted by the landlord, as long as he pays 25% of the output to the landlord

Consequences of the reform on the tenant

1 Bargaining power effect

- Tenant and landlord negotiate on the share

- Before, what would happen to the tenant if he disagreed with the landlord?

- After, what can happen to him?

- What are the consequences of this on the share of the tenant?

  - is it good or bad for productivity?

  - why?

2 Security of tenure effect

- What positive effect does it have on productivity?

-

- What negative effect does it have on productivity?

-

2.3 Empirical analysis of the reform

- Left front government came to power in 1977

- Started registration camps in villages (officials came to help tenants register)

- Faced some difficulties = flood, landlords' opposition

$\longrightarrow$ registration progressed more slowly than expected and was still continuing in early 1990s.

2.4 The expected effects of the reform

1. Reform $\rightarrow$ bargaining power $\rightarrow$ improvement in share $\rightarrow$ improvement in productivity

2. Reform $\rightarrow$ security of tenure $\rightarrow$ improvement in productivity (?)

Questions asked in the study=

a) Did reform increase share of output for the tenants?
b) Did reform increase security of tenure?
c) Did reform increase productivity?

2.5 Empirical analysis

2.5.1 Security and share of output

- From a retrospective survey: 80% said that landlord used eviction threats in the pre-reform period and 30% claimed that they (or their fathers) were actually threatened. 96% said it was difficult or impossible to evict now.

- share of tenants getting more than 50% of output went up from 17% to 39%.

2.5.2 Productivity 1

- Bangladesh

  - Neighbo ring country but no reform

- Difference in difference

	BEFORE	AFTER	DIFFERENCE
WEST BENGAL	1308.2	1649.52	
BANGLADESH	1296.76	1561.64	
DIFFERENCE	11	88	77

2.5.3 Productivity 2

- Within West Bengal

Districts had different registration rates at different times. At any given point, was productivity higher in the districts which had more registered tenants?

2.6 Regression

$$y_{dt} = \alpha_d + \lambda_\epsilon + \beta b_{dt} + \gamma X_{dt} + \epsilon_{dt}$$

α_d = district specific effect

λ = year effect

$b_{dt} =$ number of registered tenants

$X_d t =$ other district-time varying variables

$\gamma =$ effect of other district-firm varying variables on productivity

Results: [SEE TABLES 5 and 6] Show that higher registration is associated with faster growth in yield.

- Concern: Other identification issues.