Development Economics

Marcel Fafchamps
Oxford University

1. Lecture 1: Introduction and Household Models

This course is first part of the options paper in Development Economics. The second part is to take three (or more) modules among the list of 8 modules taught as part of the MSc in economics for development. The purpose of the Michaelmas course is to give students a broad perspective of the many issues covered in the development economics literature.

1.1. A brief historical overview of Development Economics

1.1.1. Development economics as a field

When development economics came to life as an economic field in the 1950's and 1960's, it had a well defined topic of enquiry: the development process. These were the days of Rosenstein-Rodan, Hirschman, Nurkse, Myrdal, and others. Development in those days was understood as more or less synonymous to growth.[1] Of course, growth is a topic of enquiry in mainstream economics. What made development economics different from the rest of economics – and necessary as a field – was the realization that growth in poor countries could not be achieved without a structural transformation of their economy from being predominantly "traditional/backward/subsistence/rural" to becoming predominantly "modern/technology-based/market-oriented/urban". This transformation process was thought to be sufficiently difficult to understand as to deserve its own label and field of economic enquiry.

Initial work on development centered around industrialization and investment. The Harod-Domar two-gap model was developed to focus attention on the need for capital and foreign exchange. The model, which is extremely simplified linear macro model, essentially says that to grow a poor country needs capital to invest and foreign exchange to purchase equipment abroad. Shortage of either will slow down growth.

Industrialization was thought in terms of synergies and clustering, very much like in economic geography or regional economics today. Hirschman developed input-output matrices as a representation of these synergies, which he called forward and backward linkages, and to guide policy. MITI – the Japanese planning body – and Asian tigers used these concepts to target specific groups of industries for growth since the 1950's. To this day, there is no agreement as to whether such targeting is successful or even useful.

Structural transformation was also seen as a labor market issue: people working in the subsistence or traditional sector would need to gradually move to cities and work in factories and service industries. In a long rambling paper, Lewis spent a few pages proposing a graphical

[1] I suspect the word development applied to poor countries goes back to the colonial era when colonial powers sought to 'develop' the new pieces of real estate they had acquired through colonial conquest. In that sense, development meant to drain the land, bring electricity and transportation, etc – do what a developer would do to increase the value of his property by building a shopping mall, an industrial estate, a residential district, or a plantation.

model of this gradual shift. This little model was to become extremely famous in spite – or perhaps because – of its simplicity. Shortly afterwards, Harris and Todaro expanded on the Lewis model by building a simple search unemployment model of rural-urban migrations. The model was used to explain a variety of things such as unemployment in cities but apparent full employment in villages, income differences between city and village, investment in education, etc. For pure historical interest, you may want to pick an old development textbook and read about the Lewis and Harris-Todaro models. You will probably be surprised by how much importance was given to these simplistic model as explaining the development process.

Over time, the requirements for structural transformation to be successful grew to include many issues that had not been anticipated by its initial proponents – such as population control, income redistribution, institutions and contracts, good governance, social capital, the welfare of women and children, protection of the environment, less crime, etc. In a 1981 article entitled "The Rise and Decline of Development Economics", Hirschman reflects on this research program, as it crumbled under its own weight and got tangled up in various policy disputes (balanced vs unbalanced growth; role of the state and industrial policy; role of trade). Perhaps because of the enormity of the task of weaving together an ever growing set of pre-requisites, development economists of that era often neglected rigorous theory and empirical work. They were rich on policy prescriptions but poor on facts. Hirschman himself – who is most known for his formal approach to policy modeling via input-output matrices – is said to have opposed too much economic modeling because he saw it as the over-mathematization of economics.

By swimming against the tide, development economics as a field ended up losing control over its original research agenda, namely how to initiate the growth process in poor countries. This research question has now been more or less captured by growth and trade economists who have recycled in their models many of the views originally proposed by development economists from the 1950's.

Unfortunately, in the process of turning over the task of understanding growth in poor countries to specialists interested primarily in rich countries, something important was lost. Familiarity with and concern for poor countries dropped because trade and growth economists typically focus on developed economies. Consequently, much of the theoretical and empirical work in trade and growth is done by people with little first hand experience of poor countries.[2] Their audience is governments in Japan, Germany, or the UK. They want to know how to get ahead of the US or how not to fall behind. They do not care (much) about the structural transformation of Kenya or Pakistan. Consequently, they typically underestimate the difficulties in initiating the structural transformation, and they tend to blame the government of poor countries for the failure of their economies to behave in the way predicted by their models.

1.1.2. Development economics as a *praxis*

Praxis is a greek word my philosophy professor used to describe the way people do things in practice. Today it is in my view futile to seek to define development economics as a field with a well defined topical question. A better, more relevant definition is: what do people calling themselves development economists do? Judging from the North-Eastern Universities Development Consortium (NEUDC) conference program,[3] pretty much anything that tickles their fancy. Development economics may have lost the territory it had initially settled. But it has

[2] Young's paper entitled A Tale of Two Cities is a good example. Yound took Singapore official data at face value and developed a whole paper to explain why it was so different from Hong Kong. It was later discovered that the reason for the difference had more to do with quirky Singapore statistics than with actual substance...

[3] The NEUDC has grown to become one of the leading development economics conference.

gained a much larger territory that it continues to explore by laying claim on more and more areas of potential enquiry. Development economics today has become synonymous with economics of poor countries. Much if not most of it is now micro. This rebirth of development economics on the ashes of the old structural transformation conundrum is for instance discussed in some of Krugman's writing, such as his papers entitled 'Towards and Counter-counterrevolution in Development Theory' and 'The Fall and Rise of Development Economics'. Bardhan has also reflected on these issues in some of his writings.

Today development economics does not differ much from other fields in terms of theory or econometric methods. So what is different between development economics and the rest of economics? Should development economics disappear as a field? I see four main remaining differences: empathy for poor countries; emphasis on fieldwork; breadth of interest; and research funding.

First of all, development economists display a concern for poor countries, a willingness to take them seriously. They naturally consider the behavior of poor peasants on the Altiplano – not just that of Wall Street traders – as worthy of economic study. I remember one famous US labor economist quoted as saying (more or less) 'Why bother study messed-up labor markets in other countries when you can study a perfect one in the US?'. To many, poor countries are indeed messed up and the policy predicament simple: become 'like us' first and then we can discuss the details. They see no point in studying institutions and markets they see as obviously inferior; for them, it is a waste of time. This approach is, for instance, reflected in editorial decisions of mainstream journals: although research in development economics covers all possible economic topics in more than three quarters of humanity, it is treated as a minor speciality. Topics that affect millions of people in poor countries are regarded as 'uninteresting to our readers'. This state of affairs is reason alone for development economics to remain as a distinct field.

Secondly, development economics emphasizes familiarity with poor countries as part of the curriculum and praxis of research, as part of sound scientific practice. This means an emphasis on field-work, surveys, and policy work, both for theorists and empirical economists. Many development economists are involved in surveys, a bit like anthropologists. This is because developing countries are much more different from each other and from rich countries than rich countries differ from each others. Familiarity with the local context and institutions are necessary to make sense of government statistics and data collected in surveys. Development economists thus spend much time visiting developing countries. To many mainstream economists, doing so is seen as a waste of time. This is because economists working on developed countries typically work on large databases collected by government agencies, partly because surveys are extremely expensive, partly because people and firms do not want to respond and the fear of interviewee fatigue restricts access to interviewees. Researchers have little influence on the type of data that are collected. Their efforts are typically geared up towards getting the data before anyone else and using them to produce top quality papers quickly.[4]

Third, development economists tend to remain less specialized than other economists. They are more likely to write one paper on, say, labor today and another on credit tomorrow. For people working on developed economies, this is nearly never true. In my view, there might be two reasons for this. First, there is a lot of demand for advice and little supply of qualified

[4] There are exceptions, of course. Some researchers collect data themselves much in the same way as development economists. In developed countries, this tradition is strongest among agricultural economists. Others generate unique datasets by combining data from multiple sources. Others yet secure access to private electronic datasets, such as bank records or ebay transactions. In all these cases, researchers have enormous difficulties circumventing confidentiality restrictions.

researchers. Consequently, researchers are often incited by funding agencies to engage in new avenues of investigation. More importantly, access to data is different. In rich countries, data sets are compartmentalized: labor economists work with worker surveys; IO economists work on enterprise data; savings specialist work on household data; etc. In development economics, household surveys are the workhorse of micro research. Much employment takes the form of self-employment: households are also farms and small businesses. Because the provision of public services is deficient, households are also where social care and public services takes place – e.g., insurance against risk, care for the sick and elderly, firewood collection, water gathering, and childcare. Having collected household data, the researcher often finds himself or herself in a position to study various aspects that in rich countries are disconnected.

Fourth, sources of research funding are also markedly different, with much funding coming from development agencies such as DfID, the World Bank, IMF, USAID, and UN agencies. Over time, their emphasis has moved away from growth, investment, and infrastructure towards poverty alleviation. Many donors feel that they cannot wait for the fruits of growth to reach the poor and that they must act now to ease their plight. Some no longer believe growth can be achieved because governments are too corrupt, inept, or both. Consequently, they do not see the point of generating growth-focused research and policy advice that will never be acted upon. Even the World Bank has joined this chorus since it now insists that all its program lending in recipient countries be based on a national poverty assessment.[5] As a result, today's development economics is characterized by much research on poverty and welfare issues (such as risk coping strategies, child labor, female health and empowerment) and on the welfare incidence of various programs (e.g., de-worming of children, micro-credit, agricultural extension to small farmers). The success of the (ambiguous) concept of human capital in mainstream economics has been used to justify many welfare interventions as investment in human capital and hence in growth. Poverty alleviation interventions are often justified as a pre-requisite for growth. While there is absolutely no doubt in my mind that, say, treating children against parasitic worms is a high priority health intervention, it is far from clear that by itself it generates high growth potential.

Ironically, development economics today seems to have lost sight of its initial objective – i.e., identifying the conditions under which poor countries can catch up with the standards of living of developed countries – and often sounds like welfare economics in poor countries. Don't get me wrong: there is nothing wrong with doing welfare economics of poor countries. In fact there is an urgent need for work in this area. My point is that development economics should recognize that studying welfare for its own sake is OK, without having to justify it as a pre-requisite for growth, but also that true, sincere work on growth is needed.

1.1.3. Recent developments

In the last few years or so, development economics has taken a few more turns. It is worth noting two relatively recent developments: (1) a dramatically increased focus on impact evaluation, randomized experiments, and the study of treatment effects more generally; and (2) less emphasis on theory and more emphasis on psychological approaches to human behavior.

The first point is taking development economics by storm, so much so that the standards

[5] This in my view is a bit strange because the World Bank lends to goverments, at concessionary rates, sure, but still must be repaid. Yet many poverty reduction or alleviation programs are welfare programs. They are like unemployment insurance or subsidized health care in developed economies, not like investment in education or infrastructure: they do not generate financial flows to serve the debt. The excessive emphasis now placed on welfare programs is something I have criticized elsewhere (see strategy report towards a growth strategy for Africa).

of proof have basically been changed – yet again (the last such revolution was the switch from cross-section to panel analysis). You'd better keep this in mind when planning your DPhil. I am teaching a month-long course on impact evaluation and treatment effects in November (part of the 2nd year MPhil Advanced Econometrics course).

The second point is related to the first and yet distinct. It is related first because experimental methods naturally lend themselves to investigation of human behavior using the tools of experimental psychology. This literature has brought to light the many ways in which human decisions fail to satisfy the rationality axiom that underlies much economic modeling – e.g., hyperbolic discounting, impulses, self-commitment problems, framing effects, rivalry in preferences and behavior, willingness to hurt oneself in order to punish/hurt others.

It is also related because impact evaluation does not easily allow the testing of structural models. This means that much empirical analysis takes the form of reduced form exclusion restriction analysis. In this context, what matters is a correct understanding of the data and the various things that can go wrong with the empirical analysis, not a correct understanding of rational agent theory. As a result, development economics has become virtually entirely empirical, with very little if any emphasis on theory, and fewer papers with even a short model to motivate the empirics. This is in sharp contrast with, say, 10 years ago.

But a focus on behavioral issues can exist quite independently from an impact evaluation/treatment effect framework. It is possible to estimate standard regressions on observational data in order to test various behavioral models. Economists are finding a lot of inspiration from behavioral, experimental, and evolutionary psychologists. Examples of ideas include:

- the altruism gene (i.e., altruism is related to genetic proximity) and its relationship with mutual insurance, the family, social networks, etc
- behavioral differences between the sexes (e.g., testosterone and aggression; pregnancy, breast-feeding and child care)
- cognitive dissonance (people hold inconsistent beliefs or fail to draw the rational implications from their beliefs), and its effect on risky behavior, e.g., casual sex
- the measurement of 'beauty' (e.g., symmetry and reversion to mean) and its relationship with nutrition and health

1.1.4. The course

When I decided to offer this course, I had to decide how to organize it. I could have started from a standard graduate development textbook, of which there are a few good ones (e.g., Debraj Ray's "Development Economics", Bardhan and Udry's textbook and readings, and Basu's "Analytical Development Economics"). I finally came to the conclusion that these textbooks do not represent the state of the field as I see it today. Also, many of the concerns of old-time development economics had already been covered in macro.

Instead, I decided to teach development economics as a praxis. I want to teach students what they would need to know to feel comfortable attending a conference such as the NEUDC. I want to cover the core of knowledge one must have in order to follow and benefit from such conferences. We therefore begin with the workhorse of development economics: the household model.

1.2. Household models

Readings: Singh, Squire and Strauss, Agricultural Household Models, The World Bank, Washington DC, 1986 – Chapters 2 and 3. An excellent discussion of household surveys and their multiple uses can by found in Deaton's book entitled "The Analysis of Household Surveys".

Most Third World producers are self-employed, as farmers or informal sector/micro-entrepreneurs. We therefore need a model that captures their behavior. The household model is nothing but a combination of the basic consumer and producer models, i.e., a consumer who also produces stuff. Early work on household models was done by Sen (JPE 1966) and Nakajima in the context of developing countries. Becker (JPE 1965) proposed a household model in the context of developed economies to understand self-provision of services such as home-cooking, child care, and the like.

We first cover unitary household models, that is, models that assume away dissension within the household as to whose preferences ought to be represented in the household's choices. If time allows, we will also discuss non-unitary household models.

1.2.1. The basic model without missing markets (as in Singh, Squire and Strauss)

The basic household model is a combination of consumer and producer models into a single model, i.e., the consumer is also the entrepreneur. We take the example of hypothetical farming household. The household has preferences on, say, an agricultural good X_a, a manufactured good X_m, and leisure X_l defined by the utility function $U(X_a, X_m, X_l)$. The household has a production technology represented by the production function $Q(L, A)$ where L is labor and A is land (acreage). Land is assumed fixed (short-run choices). The budget function of the household is:

$$p_m X_m = p_a(Q(L, \bar{A}) - X_a) + w(F - L)$$

where w is the wage rate and p_a and p_m are market prices. On the left hand side we have expenditures on manufactures. These expenditures are financed by sales of agricultural products (production minus consumption, or marketed surplus, $Q(L, \bar{A}) - X_a$) and net sales of labor on the labor market $F - L$. In addition to the budget constraint, the household must also satisfy a time accounting identity:

$$X_l + F = \bar{T}$$

where $\bar{T}$ is total time, which is fixed, and F is total family labor.

Utility maximization can be written:

$$\max_{X_a, X_m, X_l, L} U(X_a, X_m, X_l)$$

subject to the so-called full income budget constraint:

$$p_m X_m + p_a X_a + w X_l = w\bar{T} + p_a Q(L, \bar{A}) - wL$$

The left hand side of the constraint represents the value of all consumption, including own produced food and leisure. The right hand side represents the so-called full income. It is made of $w\bar{T}$, the value of total time evaluated at the market wage rate, and profits $p_a Q - wL$ (revenues minus labor costs). The full income budget constraint is obtained by replacing F using the time accounting identity.

The first order condition for optimization with respect to labor L is:

$$\lambda \left[\frac{\partial Q}{\partial L} p_a - w \right] = 0$$

where λ is the Lagrange multiplier. Since λ is always positive (more money is better, hence the budget constraint is always constraining), it can be factored out of the first order condition for labor. This implies that labor choices do not depend on λ. We are left with an expression identical to the one we would have gotten if the decision maker was a profit maximizer. Solving the above first order condition yields an optimal value of labor $L^*(p_a, w, \bar{A})$ as well as an optimal value of output $Q^*(p_a, w, \bar{A})$. If we insert it into the definition of profits, we get the profit function $\pi(p_a, w, \bar{A})$ which, combined with $w\bar{T}$, fully determines full income Y^*.

1.2.2. First order conditions and recursivity

The first order conditions for consumption goods are of the form:

$$\frac{\partial U}{\partial X_i} = \lambda p_i$$

for $i = a, m, l$, together with the full income budget constraint:

$$\sum_i p_i X_i = Y^*$$

It is easy to see that the above system of first order conditions is the same as the one we got in the consumer utility maximization problem, with Y^* replacing m. Solving the system of first order conditions yields an indirect utility function as well as the usual demand system of the form $X_i(p_a, p_m, w, Y^*)$.

The property that production decisions in such household models are independent of (i.e., can be determined without knowing) household consumption preferences is called *separability* although, strictly speaking, it should be termed recursivity. Indeed, although production choices do not depend on consumption preferences, consumption choices depend on production choices through Y^*.

1.2.3. Marketed surplus and supply response

The above household model provides a useful benchmark to study the response of market surplus $M \equiv Q - X_a$ to changes in price p_a. We begin by noting that:

$$\frac{dX_a}{dp_a} = \frac{\partial X_a}{\partial p_a} + \frac{\partial X_a}{\partial Y^*}\frac{\partial Y^*}{\partial p_a}$$

The response of consumption to a change in the price of agricultural products has two components: a standard Marshallian price effect (first term), and an income effect (second term). The income effect comes from the fact that agricultural products are one of the source of household income. Now, since $Y^* = wT + \pi(p_a, w, \bar{A})$, we have:

$$\frac{\partial Y^*}{\partial p_a} = \frac{\partial \pi(p_a, w, \bar{A})}{\partial p_a} = Q^*$$

by Hotelling's lemma. Putting everything together, we have:

$$\frac{\partial M}{\partial p_a} = \frac{\partial Q^*}{\partial p_a} - \frac{\partial X_a}{\partial p_a} - \frac{\partial X_a}{\partial Y^*}Q^*$$

We know that the first term is necessary positive (why?). The second term is normally negative (positive with negative sign in front). But the last term is normally positive (negative with negative sign in front). Consequently, the sign of the above expression is, in general, ambiguous: marketed surplus might decrease when output price increases. This will occur when the last term is large, i.e., when the income elasticity of agricultural consumption is high. To get a better sense of the anticipated magnitude of the marketed surplus response, let us rewrite the above equation in elasticity terms:

$$\varepsilon_M = \frac{\varepsilon_a}{1-s_a} - \frac{\theta_a s_a}{1-s_a} - \frac{\eta_a s_a r_a}{1-s_a}$$

where ε_a is the output price elasticity, θ_a is the consumption price elasticity, η_a is the (full) income elasticity of consumption, s_a is the share of self-consumption in output $\frac{X_a}{Q}$, and r_a is the share of agricultural revenue in full income $\frac{p_a Q}{Y*}$. With these definitions, the above equation can serve to compute the market surplus elasticity.

1.3. Household model with missing markets

Readings: de Janvry, Fafchamps and Sadoulet, "Peasant household behavior with missing markets: some paradoxes explained", Economic Journal, 101 (409): 1400-17, November 1991

1.3.1. The model

In practice, farming households in poor countries often face missing or incomplete markets, due to geographical isolation and imperfect institutions. In the presence of missing markets, households' response to market changes are likely to be quite different. The household model provides a useful framework in which to study these responses. We begin with a slightly more general notation and compact notation for the household model, borrowing heavily from de Janvry, Fafchamps, and Sadoulet (1991). The household is assume to solve the following optimization problem:

$$\max_{\{c,q\}} U(c)$$

subject the monetary budget constraint:

$$\sum_{i \in T} p_i c_i = \sum_{i \in T} p_i (q_i + K_i) + S$$

where p_i is the price of good i, c_i is the quantity consumed of good i, q_i is the quantity of good i in production (output if positive, input if negative), K_i is the household's initial endowment of good i (e.g., time), and S is an unearned supplementary cash income. The set of traded goods is denoted T. The household has at its disposal a production function:

$$G(q) = 0$$

and a self-sufficiency constraint for all non-traded goods:

$$q_i + K_i = c_i \text{ for all } i \in N$$

For notation purposes, we also note that market prices are fixed for the household:

$$p_i = \bar{p}_i \quad \text{for all } i \in T$$

This optimization problem can be rewritten as the Lagrangian:

$$\pounds : \; U(c) + \lambda \left[\sum_{i \in T} p_i (q_i + K_i - c_i) + S \right] + \phi G(q) + \sum_{j \in N} \mu_j (q_j + K_j - c_j)$$

First order conditions with respect to consumption goods are of the form:

$$\begin{aligned} U_i' &= \lambda p_i \text{ for all } i \in T \\ U_i' &= \mu_i \text{ for all } i \in N \end{aligned}$$

First order conditions with respect to production outputs and inputs are of the form:

$$\begin{aligned} G_i' &= -\frac{\lambda}{\phi} p_i \text{ for all } i \in T \\ G_i' &= -\frac{\mu_i}{\phi} \text{ for all } i \in N \end{aligned}$$

To simplify the notation, we now use a notational trick and define shadow prices for non-traded goods as:

$$p_i \equiv \frac{\mu_i}{\lambda} \text{ for all } i \in N$$

With notational convention, the first order conditions can be written as the following joint system of equations:

$$U_i' = \lambda p_i \text{ for all goods consumed} \tag{1.1}$$

$$G_i' = -\frac{\lambda}{\phi} p_i \;\text{ for all goods produced} \tag{1.2}$$

$$\sum_{i \in N,T} p_i c_i = \sum_{i \in N,T} p_i (q_i + K_i) + S \tag{1.3}$$

$$G(q) = 0 \tag{1.4}$$

$$q_i + K_i = c_i \text{ for all } i \in N \tag{1.5}$$

$$p_i = \bar{p}_i \;\text{ for all } i \in T \tag{1.6}$$

The first set of equations are the first order conditions for utility maximization; the second set of equations are profit maximization conditions; the third equation is the full income budget constraint; the fourth one is the production function (in implicit form); the fifth one is the self-sufficiency condition for non-traded goods; and the last one is the condition that prices equal market prices for traded goods. The second set of equations, together with the production function, can be used to compute the profit function $\pi(p)$. Given $\pi(p)$, full income is obtained as:

$$Y = \pi(p) + S + \sum_{i \in N,T} p_i K_i$$

Plugging the above into the budget constraint and combining with the first set of equations, we can compute the indirect utility function $V(Y, p)$. The fifth and sixth set of equations nail down the prices. The above system can thus be rewritten as a system of input demand/output supply equations, consumption demand equations, and 'market clearing' conditions:

$$q_i = q_i(p) \tag{1.7}$$

$$c_i = c_i(Y, p) \tag{1.8}$$

$$Y = \pi(p) + S + \sum_{i \in N,T} p_i K_i \tag{1.9}$$

$$q_i + K_i = c_i \text{ for all } i \in N \tag{1.10}$$

$$p_i = \bar{p}_i \text{ for all } i \in T \tag{1.11}$$

So doing, we have formally turned our household model into a little partially closed economy. The price of traded goods is determined by the market price; the shadow price of non-traded goods is determined by the self-sufficiency condition. It is as if the household had a multiple personality disorder: the producer side of the household maximizes profits, given market and shadow prices, while the consumer side of the household maximizes utility given market and shadow prices. Equations 9, 10 and 11 ensure that producer and consumer pick production and consumption plans that are consistent, i.e., that quantities produced plus initial endowments equal consumption (in case of non-traded goods), and that monies spent on the consumption of traded goods is generated by the sale of outputs or endowments.

1.3.2. Market response

The response of households with missing markets to changes in market conditions can be extremely different from that of households facing markets for all goods and factors of production. To see why, let us consider a special case with N missing markets and let us analyze what is the household's response to a change in price of an unconsumed output which, to facilitate the discussion, we call a cash crop. Totally differentiating equation 1.7 we get:

$$\frac{dq_i}{dp_i} = \frac{\partial q_i(p)}{\partial p_i} + \sum_{j \in N} \frac{\partial q_i}{\partial p_j} \frac{dp_j}{dp_i} \tag{1.12}$$

The above equation states that the effect of a change in the price of the cash crop is the sum of two effects: a direct output response and an indirect output response due to the possible effect of the change in market price on shadow prices internal to the household. To make sense of equation 1.12, we need to know how shadow prices change with market price p_i. This can be found by totally differentiating 1.10 (see de Janvry, Fafchamps and Sadoulet, 1991, for details).

Simulation exercises show that the response to a change in cash crop prices is likely to be smaller when the household is constrained to be self-sufficient in food and/or labor; it can even be perverse (why?).

1.4. Testing for separability

Optional reading: Dwayne Benjamin, Econometrica 1992

Benjamin tests whether farming households face perfect labor markets. The model he starts with is:

$$\max_{c,l,L^O,L^H,L^F} u(c,l) \text{ subject to}$$

$$\begin{aligned} c &= F(L^F + L^H, \bar{A}) - wL^H + wL^O + y \\ \bar{T} &= l + L^F + L^O \end{aligned}$$

where L^F is household labor on the farm, L^H is hired labor on the farm, L^O is household labor off-farm, l is leisure, y is unearned (exogenous) income, and c is consumption. Rearranging the budget constraint, we get the full income constraint:

$$c + wl = y + \pi(w, \bar{A}) + w\bar{T}$$

Provided that the labor market is perfect – household can sell and buy labor at wage rate w, then production decisions are separable from consumption preferences. This is because the best way for the household to maximize its utility is to maximize its income. To maximize its income, it must seek its comparative advantage and maximize profits. Since the household can buy and sell what it wants at a given price, what it wishes to consume has no effect on what it wishes to produce. This also applies to labor use: $L = L^H + L^F$ should not depend on household characteristics other than land $\bar{A}$.

Benjamin then develops some models and graphs to illustrate what happens when labor markets are imperfect and the household is constrained in its choices (show graph). In the end, the perfect market model is one in which labor is determined by:

$$F_l(L, \bar{A}) = w$$

while the imperfect market model is one in which:

$$F_l(L, \bar{A}) = w^*(\text{household composition, preferences,etc})$$

In practice, Benjamin estimates:

$$\log L = \alpha + \beta \log w + \gamma \log \bar{A} + \beta \sum_{i=1}^{G} \delta_i a_i$$

where the a_i variables stand for household characteristics such as number of household workers in various age-sex categories. Benjamin estimates this equation using data from Indonesia. Labor is for weeding, cultivation, and hoeing only. He presents various tests for robustness, including instrumenting the planting wage reported by households for hired in or out labor. In all reported regressions, number of household workers has no effect on labor use.

Problems with Benjamin's specific implementation:

1. Harvesting labour is excluded 'since it is rarely traded'. Yet this should be evidence of missing market...

2. Cultivated land is included. If the land market is very active, then households could adjust their cultivated acreage to the labor they have. Thus labor market could appear to be perfect even though it is imperfect but land adjust. This idea is particularly convincing if farming is CRS.

3. Benjamin uses household specific reported planting wage in the regression. This leads him to drop 238 observations (keeping 1443) without wage. But these are likely to be the rationed households...

In general, tests such as the one above are rejected by the data. The nice thing about this is that, if you are trying to estimate a household production function of the form:

$$Q = F(L, \bar{A})$$

you can use household endowments and preference shifters as instrument for labor: they do not affect output directly but influence the supply of labor through the household specific shadow cost of labor w^*.

1.5. Intertemporal household models

Household models can be extended to include multiple periods and risk. In the absence of risk, the intertemporal household model simply boils down to the standard neoclassical growth model. To see this, consider the following model, which is an intertemporal version of our earlier model:[6]

$$
\begin{aligned}
& \max_{\{c_t, q_t, K_t, I_t\}} \sum_{t=0}^{\infty} \beta^t U(c_{1t}, ..., c_{Nt}) \text{ subject to} \\
\sum_{i \in T} p_{it} c_{it} + p_{kt} I_t &= \sum_{i \in T} p_{it}(q_{it} + K_{it}) + S_t \\
G(q_t, K_t) &= 0 \\
q_{it} + K_{it} &= c_{it} \text{ for all } i \in N \\
K_{kt+1} &= (1 - \delta_i) K_{kt} + I_t \\
K_{it} &= K_{i0} \text{ for all } i \neq k
\end{aligned}
$$

where K_{it} denote endowment of good i (e.g., labor, capital) at time t, index k denotes a capital good, I_t stands for investment in capital, S_t are exogenous transfers, β is the household's discount factor ($\beta \leq 1$), and δ for the depreciation rate. Endowments other than capital are regarded as constant over time. The only difference with the earlier model is the addition of a law of motion for capital.

At first glance, the above model appears difficult to deal with and quite different from the standard growth model. However, it is possible to transform it into a growth model. The key is to note that, within each period, decisions about which good to consume and to produce can be solved for once we know K_t and I_t. Next we normalize prices so that $p_{kt} = 1$ for all t. The optimization model can thus be decomposed into two nested optimizations:

$$
\begin{aligned}
\max_{\{K_t\}} & \quad \max_{\{c_t, q_t | K_t\}} \sum_{t=0}^{\infty} \beta^t U(c_{1t}, ..., c_{Nt}) \text{ subject to} \\
\sum_{i \in T} p_{it} c_{it} + K_{kt+1} - (1 - \delta_i) K_{kt} &= \sum_{i \in T} p_{it}(q_{it} + K_{it}) + S_t \\
G(q_t, K_t) &= 0 \\
q_{it} + K_{it} &= c_{it} \text{ for all } i \in N
\end{aligned}
$$

If we fold investment into S_t, the inner optimization is thus exactly the same as what we have examined in the previous section and can be solved in the same way. Let total consumption expenditures be denoted $C_t = \sum_{i \in T,N} p_{it} c_{it}$. The solution to the inner optimization yields an indirect utility $V(C_t, p_t)$ and a profit function $\pi(p_t, K_t)$. As before we have $Y_t = \pi(p_t, K_t) + S_t + \sum_{i \in N,T} p_{it} K_{it} = Y(p_t, K_t, S_t)$ where the first term is pure profits, the second term is transfers from the rest of the world, and the last term is payment to factors, that is, value added. We have:

$$
\begin{aligned}
C_t &= Y_t - I_t \\
&= Y(p_t, K_t, S_t) - K_{kt+1} + (1 - \delta_i) K_{kt}
\end{aligned}
$$

[6] It is possible to obtain the same result with the full model but there is too much notation.

If there are no non-traded goods, the outer optimization model can thus be rewritten as:[7]

$$\max_{\{K_t\}} \sum_{t=0}^{\infty} \beta^t V(C_t, p_t) \text{ subject to}$$
$$C_t = Y(p_t, K_t, S_t) + (1 - \delta_i)K_{kt} - K_{kt+1} \text{ for all } t$$

which resembles the standard growth model, except for the explicit recognition that the solution depends on prices.[8] This dependence on prices is usually ignored in the notation, which is equivalent to assuming either that prices do not change over time or that output is measured in constant terms.[9] The advantage of the above presentation is to make underlying assumptions explicit.

The solution to the above is discussed in the growth lectures in 1st and 2nd year macro and need not be revisited here. Many of the issues discussed in the context of growth are relevant in the context of individual accumulation as well. For instance, everything said about poverty traps applies here as well, and many of the mechanisms discussed then arise at the micro level as well – e.g., poor by your own fault, poor because of bad circumstances, poor because of a trap, or poor because the economy is such that not everyone can be CEO or rock star. Much of what growth theory says about accumulation over time applies to individuals well, be it physical, financial, or human capital.

What was not discussed in the growth lectures is price fluctuations. There are situations in which time-vaying prices should be taken into account. Seasonal data is a case in point: relative prices are known to fluctuate systematically over the year as various crops get harvested. A seasonal analysis of intertemporal choices should therefore take price fluctuations into account. Over time, changes in household composition may also affect consumption preferences for various products – e.g., milk, nappies – and household services – e.g., childcare and elderly care. This in turn would affect $V(C_t, p_t)$ in a systematic manner. For instance, anticipated pregnancy would affect investment and labor choices of parents, especially if certain goods (milk) and services (childcare) cannot be provided by the market, possibly because of asymmetric information (milk quality, child care supervision). This, for instance, may explain why households invest less in the schooling and market-specific skills of women or why they purchase and hold cows when they have small children.

---

[7] If some markets are missing, things are even more complicated because prices of non-traded goods utimately depend on preferences and endowments. Let's partition the price vector into traded and non-traded $\{p_t^T, p_t^N\}$. We regard p_t^T as given. In general, p_t^N is a function of all variables in the model, i.e., $p_t^N(p_t^T, K_t, S_t, K_{kt+1})$. Plugging this into the indirect utility function and the profit function and definition of income gets everything in terms of capital and exogenous prices p_t^T. In this case, however, indirect utility is implicitly a function of capital as well, something that is not taken into account in the standard growth model. In practice, I have never seen an intertemporal model that takes this into account. Yet it could have important effects on savings behavior. For instance, if people with to enjoy more leisure as they get get richer, they may wish to substitute machines for labor in household production, thereby raising savings with income in a way not taken into account in the standard model.

[8] Also I have written the production function a bit differently in the sense that I expect to recover depreciated capital at the end of the period. Many growth models 'fold' this into the production function, i.e., define $f(K) = Y(K) + K$.

[9] Atanasio and Weber (JPE 1995) examine the effect of relative price changes in an intertemporal model, assuming all prices are exogenous.

1.5.1. Credit markets and separability

With a little more work, we get a result similar to the separability property of household models but this time regarding savings. Suppose capital can be rented at cost r or, alternatively, suppose that capital can be purchased with borrowed funds, whereby the interest on borrowed funds is r. In this case, the optimization problem becomes (dropping some of the notation to improve readability):

$$\max_{\{K_t,W_t\}} \sum_{t=0}^{\infty} \beta^t V(C_t) \text{ subject to}$$
$$C_t = Y(K_t) - \delta_i K_t - rK_t - W_{t+1} + (1-r)W_t$$

where W_t is the net wealth of the household, and r is the interest rate. I have assumed that the household can save and earn an interest r. Since only K_t enters the budget constraint, not K_{t+1}, the Euler equation for K simply is in each period:

$$\frac{\partial Y(K_t)}{\partial K_t} = r + \delta$$

which means that the marginal return to capital equals the rental cost of capital plus depreciation. In this case, production decisions are independent from endowments (e.g., wealth W_t) and preferences (e.g., discount factor β). The producer adjusts instantaneously to the optimal capital level. There is no waiting time.

There is a lot of work in development economics focusing on credit constraints and the need for the poor – and poor countries in general – to accumulate before being able to make a large, lumpy profitable investments. Nurkse (Problems of Capital Formation in Underdeveloped Countries, 1953) made the point that capital formation is the key constraint of poor countries. Gerschenkron (Economic Backwardness in Historical Perspective, 1962) went further by insisting that providing funds to the country is not enough; it must reach investors. He started a large literature on the role of financial intermediation in the development process. McKinnon (1973) further argued that to speed up investment by the poor, savings instruments available to them must yield a high enough return.

Much of the literature on poverty has concerned itself with this issue – e.g., current efforts to develop micro-credit. Two observations are in order. First, the absence of credit does not, by itself, preclude social mobility; it only slows it down: in the model presented above, poor people save and eventually get out of poverty. Second, this model does not include any risk. If we include risk, then giving credit to the poor may, if they are unlucky with their investment, push them into a debt trap which, in the absence of personal bankruptcy law, can last indefinitely. To put differently, contrary to the way many people think, there is no poverty trap in the absence of credit – just a delay. But there is a debt-driven poverty trap once credit to the poor is introduced and some risk exists. Put differently, without credit a poverty trap does not exist but with credit one appears... This is an issue I discuss in my paper with Gubert on contingent loan repayment.

1.6. Investment and poverty trap

Equipped with the above theoretical apparatus, we are in a better position to discuss various explanations that have been offered for poverty traps and poverty persistence. Suppose that there is an investment opportunity with production function $Y(K)$. Variable K is not limited to capital and could be anything – purchase of capital equipment, investment in human capital, adoption of new technology, effort to build social contacts. Instantaneous utility of consumption is $V(C_t)$ and the decision maker discounts the future with discount factor β, as before. We assume that $Y'(K) >> 1/\beta - 1$ for some K, i.e., the investment is very profitable.

A poverty trap arises if the decision maker cannot invest in K. Poverty persistence arises if it takes a very long time for the decision maker to invest in K. In the short term, the two concepts are observationaly very similar and the literature often uses them interchangeably. There is no poverty trap if the agent invests in K. Under what condition does the agent invest in K?

We first note that if credit markets are perfect – or more generally if the producer can rent all the productive equipment – the decision model can be written as above:

$$\max_{\{K_t,W_t\}} \sum_{t=0}^{\infty} \beta^t V(C_t) \text{ subject to}$$
$$C_t = Y(K_t) - rK_t - W_{t+1} + (1-r)W_t$$

where for simplicity I have subsumed depreciation in the production function $Y(K)$. As shown above, the Euler equation for K simply is in each period:

$$Y'(K) = r$$

that is, marginal return equal marginal cost.

If there is no credit market, the optimization problem becomes:

$$\max_{\{K_t,W_t\}} \sum_{t=0}^{\infty} \beta^t V(C_t) \text{ subject to}$$
$$C_t = Y(K_t) + K_t - K_{t+1}$$

[You will note that this is very similar to a neo-classical growth model.] The Euler equations for K simply is in each period:

$$Y'(K) = \frac{1}{\beta}\frac{V'_{t-1}}{V'_t} - 1$$

If the return to the investment is large, C_t will be larger than C_{t-1}. This implies that $V'_{t-1} > V'_t$ – the marginal utility of income in the future is lower than today. It follows that, even if $1/\beta \approx 1 + r$ (as should be the case in a perfect equilibrium), we get:

$$Y'(K) > r$$

which means that there is underinvestment in K. Note that when capital is divisible, there is not strictly speaking a poverty trap: the agent underinvests but the high return to K at low levels of investment incite him/her to invest anyway. Over time, the agent should be able to undertake the investment. In other words, there is poverty persistence but not a strict poverty trap. [King and Rebelo presents simulations of this model, which is also the neo-classical growth model.]

For a poverty trap to exist, the investment must be non-divisible. For instance, if the return to K is negligible below a threshold $\overline{K}$ but high at $\overline{K}$, a poor individual must first accummulate $\overline{K}$ using personal savings at a low or negligible return. This may take a long time. If the time taken is too long, it may be optimal for the decision maker to forget about the investment all together. In a situation like this one, it is easier for the poor investor to invest if reliable savings instrument exist and if the return on savings is large, an argument initially made by McKinnon in 1973. In many countries this is not the case: poor household have few or no reliable savings instrument in banks and financial institutions, and cash savings often have a negative return because of inflation. Inflation can thus be a cause of poverty trap.

The missing market/separability logic can be extended to other types of market failures. For instance, suppose that credit markets exist but there is insufficient protection of property rights and no insurance can be purchased against the risk of expropriation. The optimization problem becomes:

$$\max_{\{K_t,W_t\}} \sum_{t=0}^{\infty} \beta^t V(C_t) \text{ subject to}$$
$$C_t = (1-\omega)Y(K_t) - rK_t - W_{t+1} + (1-r)W_t$$

where $\omega \leq 1$ is the probability of expropriation. If the decision maker is risk neutral, the Euler equations for K now is:

$$Y'(K) = \frac{r}{1-\omega} > r$$

which shows that the risk of expropriation reduces the expected return from investment – and hence investment itself. Here the missing market is insurance against expropriation.

It is also possible to obtain a similar result for investment under uncertainty when the decision maker is risk averse. As shown by Sandmo (1970), the absence of insurance market leads to underinvestment relative to risk neutral/perfect insurance market case. Poor households customarily face a lot of risk, especially in rural areas (see my book on Rural Poverty, Risk and Development). They are also imperfectly insured, as a voluminous literature has shown. Consequently, the literature concludes, it is not surprising that they underinvest. (See however my recent paper entitled "Vulnerability, Risk Management, and Agricultural Development" where I draw on the recent experimental evidence to challenge some of these ideas.)

1.6.1. A special case: human capital

We focus here on the schooling dimension of human capital. There is an enormous literature on this. Here I develop a little model and illustrate how it can be used to derive interesting results.

Being just another form of capital, human capital can in principle be modeled in exactly the same way as capital. There is a price to pay – school fees, foregone income – and a return – increased wages or earnings from self-employment. Let the cost be written $F(E_{t+1})$ and the wage function be written $w(E_t)$ with $F' > 0$ and $w' > 0$. Let us ignore other assets – and assume there is no credit market. Furthermore, for simplicity, consider an individual who lives two periods – works or goes to school in period 1 and only works in period 2. We have:

$$\max_{\{E_2\}} V(C_1) + \beta V(C_2) \text{ subject to}$$
$$C_1 = w(0) - F(E_2)$$
$$C_2 = w(E_2)$$

which can be simplified into:

$$\max_{\{E_2\}} V(w(0) - F(E_2)) + \beta V(w(E_2))$$

with first order condition:

$$-V_1'F' + \beta V_2'w' = 0 \text{ or}$$
$$w' = F'\frac{V_1'}{\beta V_2'}$$

which means that the marginal return to an additional year of schooling is equated with the marginal cost of acquiring education times $\frac{V_1'}{\beta V_2'}$. The latter term is a correction for the difference in levels of consumption between the two periods: in general we assume that V' falls with consumption – the marginal satisfaction one gets from an additional unit of consumption gets smaller as income rises. If this is true, $V_1' > V_2'$ if $C_1 < C_2$. Since $w(0) < w(E)$ in general we have $V_1' > V_2'$ which implies that $w' > F'$ in most cases. Put differently, individuals underinvest in education because of the high utility cost of paying school fees when consumption is low.

As before, underinvestment disappears if we have a perfect credit market. To see this, add debt D to the model. We now have:

$$\max_{\{E_2,D\}} V(C_1) + \beta V(C_2) \text{ subject to}$$
$$\begin{aligned} C_1 &= w(0) - F(E_2) + D \\ C_2 &= w(E_2) - (1+r)D \end{aligned}$$

Solving for the two first order conditions, we obtain:

$$\begin{aligned} -V_1'F' + \beta V_2' w' &= 0 \\ V_1' - (1+r)\beta V_2' &= 0 \end{aligned}$$

from which we get:

$$\begin{aligned} \frac{V_1'}{\beta V_2'} &= 1+r \\ \frac{w'}{1+r} &= F' \end{aligned}$$

which shows that the (discounted) marginal return to education is set equal to the marginal cost.

These simple observations form the basis of many efforts to help the poor get education – e.g., give student loans, reduce the cost of going to school (free tuition, free meals, etc) or, possibly, make it illegal for parents to send their kids to work (eliminating the opportunity cost of foregone child labor).

It may be of interest to note that this model suggests that the best method is student loans, not low tuition. To see this, suppose that households differ in initial endowments, so that some people can consume some extra amount W_0. In this case, it is impossible to find a tuition grant that ensures that $\frac{w'}{1+r} = F'$ for all households. In fact, a tuition grant equal for everyone would raise education demand by the 'privileged' (lower V_1') so that, in the end, they would get more education than in the first best equilibrium – and than the poor. Of course, starting from a situation in which education is free, the institution of a tuition fee would represent an immediate financial loss that is felt more by poor parents, given that they are less able to absorb it from their accumulated savings. Showing this formally is left as an exercise.

1.6.2. Human capital as bequest

The above human capital model assumes that the investment is done by the person herself. In most cases, however, education is financed by parents as a form of bequest inter vivos.[10]

[10] This need not be true for university education, but it is true for primary and secondary education.

Schooling then is the result of an altruistic decision by parents. One way of modeling this is to write a modified human capital investment model as:

$$\max_{\{E_2\}} V_1(C_1) + \omega V_2(C_2) \text{ subject to}$$
$$\begin{aligned} C_1 &= w(\overline{E}_1) - F(E_2) \\ C_2 &= w(E_2) \end{aligned}$$

where for simplicity I have assumed that parents live only in period one while their child studies in period 1 and works in period 2. Parameter ω captures parental altruism, that is, how parents value the utility of their child $V_2(.)$. It is immediate to see that the solution to the above optimization problem is very similar:

$$w' = F' \frac{V_1'}{\omega V_2'}$$

If parents value their child's utility at most as much as their own, and they are poorly educated with a low income, then we have the same result as before: $w' > F'$, meaning that there is underinvestment in the child's education. If, however, $\omega >> 1$, it is conceivable that parents will invest at the optimal level – or even over-invest – although they are poor.

More complex models can be investigated in which parents can choose whether to bequeath their wealth in terms of human capital or real assets, e.g., land. Quisumbing (1994) for instance shows that rural Filipino parents educate their daughters but bequeath land to their sons.

As a final note, we have presented the modeling framework as a model of investment in education. But it could just as well be interpreted as a model of investment in nutrition or health – or any other form of human capital.

1.7. Applications and empirical work

Apart from the massive influence these models have had on the thinking about poverty traps and poverty persistence, there is an enormous body of empirical work that uses household models as starting point. In fact, virtually all microeconomic work on rural households uses this modeling framework in one way or another. Here are just a few examples:

- Test of separability of the household model (e.g., Benjamin, Econometrica, 1992).
- Estimation of separable household models to compute marketed surplus supply response (e.g., Singh, Squire and Strauss, 1986).
- Estimation of (parts of) non-separable household models:
  - output supply;
  - input demand;
  - consumption demand;
  - labor supply;
  - investment;
  - borrowing;
  - consumption of/investment in welfare type goods such as education, health, child nutrition, micro-nutrients, etc.

In many cases reliance on household models is only implicit. For instance, it is common for researchers to estimate a production function using village-level data. But only if certain markets are incomplete/imperfect is the estimation possible. Indeed, if all producers faced the same factor prices and had the same production function, they would use the same level of inputs. This would make estimation impossible. Hence the estimation of production functions implicitly assumes non-separability/missing markets.

2. Lectures 2-3: Market institutions and imperfect commitment

Readings:

Platteau, J.P., "Behind the Market Stage Where Real Societies Exist: Part I - The Role of Public and Private Order Institutions", *Journal of Development Studies*, 30(3): 533-577, April 1994

Platteau, J.P., "Behind the Market Stage Where Real Societies Exist: Part II - The Role of Moral Norms", *Journal of Development Studies*, 30(4): 753-815, July 1994

Greif, A., "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition", *American Economic Review*, 83(3): 525-548, June 1993

Bigsten, A. et al., "Contract Flexibility and Dispute Resolution in African Manufacturing", *Journal of Development Studies*, 36(4): 1-37, April 2000

Coate S., and G.C. Loury, "Will Affirmative Action Policies Eliminate Negative Stereotypes?", *American Economic Review*, 83(5): 1220-1240, December 1993

Fafchamps, M (2004), *Market Institutions in Sub-Saharan Africa: Theory and Evidence*, MIT Press, Cambridge, Mass.

Fafchamps, M., "The Enforcement of Commercial Contracts in Ghana", *World Development*, 24(3): 427-448, March 1996

Fafchamps, M., "Spontaneous Market Emergence", *Topics in Theoretical Economics*, 2(1): article 2, 2002

Fafchamps, M., "Ethnicity and Credit in African Manufacturing", *Journal of Development Economics*, 61(1): 205-235, February 2000

Fafchamps, M., (2003) "Ethnicity and Networks in African Trade", Contributions to Economic Analysis and Policy, 2(1): article 14, Berkeley Electronic Press at www.bepress.com

2.1. Market institutions and development

Since North's seminal work on the role of good market institutions in the historical development of Europe, it is widely recognized that market institutions are a critical pre-requisite for development. The purpose of this lecture is to present a consistent framework with which to approach market institutions in general. Evidence from my 2004 book on market institutions in Africa will be provided in class, time permitting.

Exchange of goods and services between agents is essential for economic development. This is because exchange makes specialization possible, and specialization is essential for economic growth. Without exchange, people have to be self-sufficient. This means providing to themselves much of what they consume – food, housing, simple tools and toys, entertainment, insurance, health care, personal grooming, child care, elderly care, etc.

Households living in poor undeveloped parts of the world still live largely in autarky. This means that they are very unspecialized. Lack of specialization means that the scale of production is minuscule, hence making it impossible to capture returns to scale. Because people undertake so many different tasks, they have little time and inclination to learn much about how to improve productivity in any single one of them. As a result their productivity is low – essentially they produce crude goods and procure low quality services. Without specialization, it is difficult if not impossible to capture the productivity gains brought about by the application of science to technology: there is little learning and because of the small scale of production there is little opportunity to take advantage of technological change embedded in large specialized equipment.

2.1.1. Modes of exchange

There are three main modes of exchange: gift exchange; market exchange; and exchange within hierarchical organizations. Gift exchange is what takes place, for instance, within households: one person ploughs the land, another fetches the water. Because gift exchange is based either on altruism or anticipated reciprocity, its effective 'range' is very narrow: friends and family. It is difficult to rely on gift exchange for sizeable exchange beyong these narrow confines. Increased specialization normally means that the range of exchange covered by gift exchange shrinks – e.g., people no longer cook their own food, they go to the restaurant instead. This means that specialization brings out the role of the market.

It is also possible to organize exchange through command and control within a hierarchical organization such as a firm, a government ministry, or an NGO. Communist countries basically tried to organize almost all economic exchange through command and control. This proved difficult. But we should not forget that the 20th century has seen the rise of public and private organizations of a size unheard of before the industrial revolution. For many developing countries, setting up these large organizations in an effective manner has proved a challenge, as the literature on corruption, governance, and ineffective government has documented. We do not have the time to discuss why, but suffice to say that if gift exchange cannot accommodate the needs of an advanced economy with a high degree of specialization, and large organizations are difficult to set and workers/civil servants difficult to discipline, then this leaves a big role for the market. This is basically what has been taking place in Africa over the last century or so: the replacement of gift exchange by market exchange and the massive rise of private enterprise, albeit in the form of a myriad of very small enterprises. This means that in developing economies market exchange is probably responsible for a larger share of total exchange than in very primitive economies (where gift exchange still looms large) and advanced economies (where much exchange takes place within large organizations).

It follows that market institutions are essential for economies at early stages of development. (They are also essential for more advanced economies where exchange is critical in highly specialized goods and services – think of industrial services, for instance, such as warehousing, transport, advertizing, accounting, finance, insurance, cleaning services, product design, exporting and importing agents, internet design, maintenance, custodial services).

2.1.2. Laws and courts

It is customary for economists to think that (all) what is needed for market institutions are laws and courts – and that the economist's role is simply to ensure that courts are funded and basic laws are instituted that protect property rights. This view is too simplistic.

For one thing, in some fundamental sense, courts can never *enforce* a contract:[11] if a contractual dispute ends up in court, it is because the contract has not been respected. All the court can do, after the facts, is to provide reparation by awarding damages, etc. In other words, the court seeks to correct for the fact that the contract has not been enforced. Only the *threat* of court action can enforce a contract. But for this threat to have a deterrence effect, it must be credible.

As it turns out, there may be other forces that have a deterrence effect on opportunistic breach of contract. One cannot understand the role of laws and courts in enforcing contracts without a proper grounding in the origin of (rational) trust. This is the object of this lecture.

[11] I have sometimes heard economist draw a distinction between market exchange and contracts. This is of course non-sensical. ANY market transaction is a contract – a sales contract, loan contract, employment contract, rental contract, etc.

2.2. Trust and trade

Market institutions are a critical and little understood component of a good business environment. When foreign firms invest in poor countries, they often are surprised that contracts are not fulfilled as anticipated. The 'business culture' as some put it appears to be different. This generates quite a lot of animosity and, probably, discrimination as well. There is a general perception that poor countries suffer from a 'commitment problem' – i.e., governments and people cannot keep their promises – and poor countries are often thought to be unsafe and risky places to invest for this reason. It is therefore important that we seek to understand how markets operate in such countries.

2.2.1. Mechanisms for contract enforcement

For trade to take place, economic agents have to find someone they can trust to conduct a bona fide transaction. There are many opportunities for cheating in trade, from misrepresenting quality to absconding with payment. This raises the issue of trust between potential buyer and seller.

Economics suggests a way of thinking about trust, not as an emotion, but as a rational thought process. To trust someone 'rationally', we must believe that this person has adequate incentives to behave in a trustworthy manner. We define incentives in a broad manner. Several incentives have been suggested in the literature: guilt; the fear of legal reprisal; the fear of strong-armed enforcement; the fear of spoiling a long-term profitable relationships; and the fear of losing one's reputation with others.

Guilt is internal to each individual. One's ability to feel guilty for failing to respect business promises varies among individuals. Honesty is largely the by-product of upbringing, what psychologists call 'secondary socialization'. It is also influenced by cultural values and religious beliefs. Enforcement mechanisms that rely on coercion are of two types: legitimate and illegitimate. The legal enforcement of contracts through courts ultimately relies on the state's monopoly over legitimate force. It is the state's backing that allows buyers to seize a debtor's assets and thus grants collateral value to unmovable property. Illegitimate force can also be used to enforce contractual obligations. Parties may resort to insults and violence directly, or hire thugs and bribe policemen to intervene. In the great majority of cases, the actual use of force is not required; implicit or explicit threats are sufficient.

Threats, however, are not always credible. Indeed, whether legitimate or illegitimate, the use of coercion to enforce contracts is costly. For small transactions, legal costs are typically too high to justify court action. Whenever the threat of coercion is not believable, it fails to induce compliance unless the offending party can be persuaded the aggrieved party will go to court or resort to violence even at a cost – e.g., because she wishes to preserve a reputation of toughness, because she is angry, or because her moral sense compels her to do so.

The third type of enforcement mechanism is based on quid pro quo: 'I continue to behave if you continue to behave'. It is the threat of retaliation that induces compliance with contractual obligations. For such a mechanism to work, parties must interact repeatedly over time. The simplest form of retaliation is the refusal to further transact. For this punishment to deter breach, the relationship must be something worth preserving. Retaliation may also be inflicted by a group of people who were not party to the contract. Any group punishment requires a coordination mechanism and the circulation of information about contract compliance within the group. Reputation is that coordination and information sharing device. Enforcement mechanism based on reputation are vulnerable to disinformation: they do not operate well unless a complementary mechanism ensures the accuracy and veracity of the shared information.

2.2.2. A formal model

These concepts are now illustrated formally. Consider a contract by which a buyer promises to pay f at time 1 to a seller in exchange for a quantity k at time 0. Other contractual obligations (delivery on time, warranty, quality, etc) can be analyzed in a similar manner.

The set of subgame perfect contracting equilibria – i.e., of contractual promises backed by credible threats – is derived by backward induction. At time 1, the buyer decides whether or not to comply with the contract. The cost of complying in general varies with the buyer's type and unanticipated shocks. For instance, a financially sound trader (good type) will find it easier to pay on time than an over-leveraged trader (bad type). Similarly, a buyer may not pay because of cash flow problems (bad shock). The cost to the buyer of paying f can thus be written as $\pi(f, \tau, \varepsilon)$ where τ denotes the buyer's type and ε denotes the state of nature at time 1. Type $\tau \in \Delta$ is any characteristic of the buyer that is relevant to the contracting situation, like his or her professional experience, technology, preferences, and honesty. The state of nature $\varepsilon \in \Sigma$ is any condition exogenous to the parties that was unknown at time 0 and makes payment harder or easier. If compliance is totally impossible, we say that $\pi(f, \tau, \varepsilon) = \infty$. How severely shocks affect buyers' ability to fulfill the contract in general depends on their type: those who are less competent or ill prepared have a high cost of compliance $\pi(f, \tau, \varepsilon)$ for many states of nature. The function $\pi(f, \tau, \varepsilon)$ allows for these effects as well. We assume that the sets of possible types Δ and states of nature Σ are common knowledge, but that only the buyer knows his or her type τ. Some shocks are observable ex post by both parties; others are known only to the buyer. As a result, the buyer has better information about ε than the seller.

In case of non-payment, the buyer receives a payoff of 0 but incurs punishment. We consider four types of punishments that correspond to the three categories discussed above: guilt, whose utility cost to the buyer is denoted $G(\tau, \varepsilon)$; various forms of coercive action including harassment, threats, and court action, whose cost to the buyer is denoted $P(\tau, \varepsilon, C)$; and two types of punishments based on repeated interaction: the suspension of future trade with the seller resulting in the loss $EV(\varepsilon, \tau)$; and damage to the buyer's reputation with other potential trading partners leading to a loss $EW(\varepsilon, \tau)$. The term $EV(\varepsilon, \tau)$ represents the value of the relationship, that is, the expected discounted value of future transactions with the seller; $EW(\varepsilon, \tau)$ is the value of lost reputation, that is, the expected discounted value of future transactions with all those who will refuse to transact with the buyer after a breach has occurred. A rational buyer fulfills the contract if the cost of complying is smaller than all penalties combined, i.e., if:

$$\pi(f, \tau, \varepsilon) \leq G(\tau, \varepsilon) + P(\tau, \varepsilon, C) + EV(\tau, \varepsilon) + EW(\tau, \varepsilon) \tag{2.1}$$

Whenever $\pi(f, \tau, \varepsilon) = \infty$ the buyer is unable to pay and the contract is breached. There are also situations in which the buyer could in theory pay, that is, $\pi(f, \tau, \varepsilon) < \infty$ but equation (2.1) is not satisfied. The buyer is then said able but unwilling to pay.[12] By definition, a buyer who is unable to pay is also unwilling to pay. Penalties in general depend on the buyer's type τ and on the realized state of nature ε. For instance, some agents are unscrupulous and have a low $G(\tau, \varepsilon)$. Others are hard to harass and coerce into paying their debts through legal (or illegal) means and have a low $P(\tau, \varepsilon, C)$. Others yet, like fly-by-night operators or firms on the verge of bankruptcy, have a short horizon and little interest in preserving their reputation – low $EW(\varepsilon, \tau)$– and their relationship with the seller – low $EV(\varepsilon, \tau)$. All these effects are accounted for in equation (2.1). The strength of harassment and threats also depends on the form of the

[12] The distinction between inability and unwillingness to repay is blurred in practice. For equity reasons, debtors often are regarded as unable to repay when compliance would be unduly costly, i.e., when $\pi(-f, \tau, \varepsilon)$ falls below a socially unacceptable level $B < \infty$.

contract C, e.g. on whether formal guarantees were provided or whether contractual obligations were put down in writing to ease the burden of proof.

2.2.3. The seller's incentives

Now consider time 0. The seller is asked to part with k in exchange for a future promise of f. Let $\Pi(k)$ and $\Pi(f)$ be the value of k and f to the seller. We assume that there are gains from trade: $\Pi(k) < \Pi(f)$. In forming beliefs about the likelihood of receiving f, a rational seller evaluates the chances of being paid, i.e., the probability that equation (2.1) will be satisfied. In evaluating this probability, the seller uses all the information, denoted Ω, available at time 0: prior knowledge about the distribution of potential buyer types, information gathered over time through direct interaction with the buyer, and information conveyed by others about the buyer. Formally, let $F(\tau, \varepsilon|\Omega)$ be the joint cumulative distribution over τ and ε that captures the seller's beliefs given information Ω. Rank states of the world so that, for any buyer type τ, $\pi(f, \tau, \varepsilon)$ is decreasing in ε: it is easier to comply in good states. Further assume that each of the four penalties listed in equation (2.1) is non-decreasing in ε, i.e., that the buyer has more to lose in good than in bad states. We can then define the function $h(\tau)$ as the level of shock ε at which equation (2.1) is exactly satisfied and a buyer of type τ is just indifferent between compliance and breach, i.e. $h(\tau) = \varepsilon^*$ such that:

$$\pi(f, \tau, \varepsilon^*) = G(\tau, \varepsilon^*) + P(\tau, \varepsilon^*, C) + EV(\tau, \varepsilon^*) + EW(\tau, \varepsilon^*) \tag{2.2}$$

For notational simplicity, let us ignore the possibility of partial payment. Then, for any shock ε above $h(\tau)$ the buyer pays; for any shock below $h(\tau)$ no payment is made. Let $(\underline{\tau}, \bar{\tau})$ and $(\underline{\varepsilon}, \bar{\varepsilon})$ be the lowest and highest values that τ and ε can take. A rational seller agrees to a contract (k, f) if and only if what he or she expects to receive is greater than what is given, i.e., if:

$$\begin{aligned} E(\Pi(f)|\Omega) &= \Pi(f) \Pr(payment) && (2.3) \\ &= \Pi(f) \int_{\underline{\tau}}^{\bar{\tau}} \int_{h(\tau)}^{\bar{\varepsilon}} dF(\tau, \varepsilon|\Omega) \geq \Pi(k) && (2.4) \end{aligned}$$

Equation (2.3) can be understood as follows. If the buyer's type were known to be, say, τ, the probability of being paid would be equal to the probability that the exogenous shock ε is greater than $h(\tau)$, i.e., to $\int_{h(\tau')}^{\bar{\varepsilon}} dF(\varepsilon, \tau|\Omega)$. Since the seller does not know the buyer's type, the probability of being paid must be computed over all possible types, hence the double integral in equation (2.3).

The seller may be able to affect the probability of repayment by adjusting the form C of the contract. For instance, the seller may request that the buyer mortgages real assets to service the debt in case the buyer goes bankrupt. Arranging legal security is costly and time consuming, however. Say there are N possible forms the contract C_n can take, each with its own cost B_n. The seller then must choose a contractual form C_n such that the value of the transaction net of transaction cost $E(\Pi(f)|\Omega) - \Pi(k) - B_n$ is maximized. The solution to this optimization problem may be to bypass formal guarantees if contract enforcement mechanisms other than $P(\tau, \varepsilon, C)$ are sufficient. If commercial transactions can be enforced through repeated interaction alone, i.e., through $EV(\varepsilon, \tau)$ and $EW(\varepsilon, \tau)$, one expects them to make little or no use of formal guarantees and of the court system. Similarly, one expects legal institutions providing a lot of security at a high cost B_n to be most relevant for large anonymous transactions (such as the sale of a house).

There may also be situations in which, for all possible contractual forms C_n the net value of the transaction is negative. In those cases, the seller refuses to trade. Imperfect enforcement then results in rationing. This is one of the possible equilibrium configurations resulting from hidden information models. Small transactions, for instance, are difficult to enforce through courts and, if they are anonymous, cannot rely much on expected future trade. As a result, one expects small anonymous transactions to be self-liquidating, with immediate cash payment and no delayed obligations. This form of trade, which I have called 'flea market economy' in one of my papers, characterizes much of the observed trade in agricultural output. Needless to say, it is not a very efficient way of proceeding as it raises the risk of theft and requires that transactions be conducted in person – not by telephone or fax. This in turn raises transport costs (for the trader) and limits the size of firms (trader too busy running around from market to market).

2.2.4. The buyer's ex ante incentives

The buyer also must agree with the contract *ex ante*. A rational buyer will do so if and only if he or she expects to derive a benefit from the contract. The buyer knows his or her type, say, τ'. Let then $\pi(k,\tau')$ denote the value of receiving k for the buyer and, again for notational simplicity, ignore partial payments. In period 1, either the buyer pays and incurs a cost $\pi(f,\tau',\varepsilon)$, or does not pay and incurs the punishments listed in equation (2.1). Given the buyer's type, payment occurs with probability $\int_{h(\tau')}^{\bar{\varepsilon}} dF(\varepsilon|\tau')$. The buyer therefore agrees to the contract if and only if:

$$\pi(k,\tau\prime) \geq \int_{h(\tau')}^{\bar{\varepsilon}} \pi(f,\tau',\varepsilon)dF(\varepsilon|\tau\prime) + \tag{2.5}$$

$$\int_{\underline{\varepsilon}}^{h(\tau')} [G(\tau\prime,\varepsilon)+P(\tau\prime,\varepsilon,C)+EV(\tau\prime,\varepsilon)+EW(\tau\prime,\varepsilon)]dF(\varepsilon|\tau\prime) \tag{2.6}$$

Equation (2.5) states that the buyer's gain from the contract (first term) must be greater than the expected cost of complying when compliance occurs (second term) plus the expected cost of punishment when compliance does not occur (third term).

Equations (2.3) and (2.5) illustrate the tension inherent to any contract. If enforcement is too lenient, buyers will promise anything knowing that if they breach their promise, they will not be penalized. At the limit, if enforcement is zero, $h(\tau)=\bar{\varepsilon}$, the seller expects no payment at all and no contract is concluded. Similarly, if enforcement is very harsh, say infinite, and even the best buyers are occasionally unable to comply, then the expected cost of punishment ∞ is larger than the gain from any contract. As a result, the buyer refuses to promise something he or she is not sure to deliver. In both cases, no contract is concluded even though there may be significant gains from trade. For trade to occur, enforcement must be sufficiently strong to deter opportunistic breaches but not so strong that it scares away all potential buyers.

A finite but high penalty for non-payment may be used to induce good types (e.g., with sufficient financial backup) to reveal themselves by buying on credit while bad types (e.g., little financial backup, high risk of financial difficulties) are induced to trade cash. In the case of agricultural trade, legal sanctions are seldom sufficient to deter bad types from buying on credit. This is because bad types are poor with little or nothing to foreclose on. Moreover, the size of the transactions is often too small to justify going to court. In these cases, the threat of court action is not credible. Other incentives must be relied upon.

2.3. The value of a relationship

The set-up outlined above is fine as far as it goes but it is not specific enough. What is the value of the relationship $EV(\tau, \varepsilon^*)$? What is the value of reputation $EW(\tau, \varepsilon^*)$? To these we now turn. We begin with $EV(\tau, \varepsilon^*)$. As will become rapidly clear, the value of $EV(\tau, \varepsilon^*)$ depends not only on the characteristics of the transaction but also on the trade environment.

2.3.1. Two agent example

To illustrate the value of a relationship, I begin with the following simple example. Consider two agents, a client A and a supplier B. For ease of presentation, we assume that they are in a long-term relationship with no end in sight. Each month, A receives merchandises from B and must pay upon receipt of a monthly invoice. If A pays, his gain is his profit margin α. If A does not pay, his gain is the value of the good – normalized to be 1 – and the profit margin: $1 + \alpha$. The client has a discount factor:

$$\beta = \frac{1}{1+\delta} < 1$$

where δ is the (monthly) rate at which A discounts the future. If A pays, the relationship continues and more supplies arrive the following month. If A cheats, B stops supplying forever.

We are now in a position to compute the value for A of the relationship with supplier B. Consider what happens in case of breach of contract. If A does not pay, he gains 1 this month but B will refuse to trade from then on. Consequently, A loses the value of α for all future periods, that is, $\sum_{t=1}^{\infty} \beta^t \alpha$ which is equal to $\frac{\beta\alpha}{1-\beta}$. In this case, the client would choose not to cheat if:

$$1 + \alpha + \beta 0 \leq \alpha + \frac{\beta\alpha}{1-\beta} \tag{2.7}$$

$$1 \leq \frac{\beta\alpha}{1-\beta} \tag{2.8}$$

Inequality 2.7 is called the voluntary participation constraint or non-cheating constraint. Provided β is close enough to 1 (that is, provided the discount rate δ is small enough) and α is strictly positive, $\frac{\beta\alpha}{1-\beta}$ can be arbitrarily large. To facilitate comparison with the first section, note that inequality 2.7 can also be derived by in terms the instantaneous gain from cheating $1 + \alpha - \alpha = 1$ and the long-term loss from losing the relationship:

$$\pi(f, \tau, \varepsilon) = 1 \leq \frac{\beta\alpha}{1-\beta} = EV(\tau, \varepsilon^*) \tag{2.9}$$

We see that the future value of the relationship is simply $\frac{\beta\alpha}{1-\beta}$. For β close enough to 1, the fear of losing the relationship can, by itself, ensure respect of the contract.

2.3.2. N-agent example

The above example illustrates the value of a relationship if no outside option exists at all. We now examine what happens if outside options exist. We imagine a situation with two groups of agents, clients and suppliers, trading repeatedly over time. As before, each transaction is such that payment takes place after delivery, so that there is a contract enforcement problem. There are two types of suppliers: good and bad. Good suppliers deliver quality stuff, bad suppliers do not. When the client buys from a bad supplier, he makes zero profit. When he buys from a good supplier, he makes α profit as before.

The type of an individual supplier is not immediately observable. To discover the supplier's type, the client has to experiment, i.e., purchase a sample and try it out. The cost of experimentation is $c > 0$. The proportion of good suppliers in the total population is θ.

A client without supplier has to sample a supplier at random and incur cost c to find out whether the supplier is good or bad. If the supplier turns out to be good, they enter in a long-term relationship in the following month. If the supplier is bad, the client has to sample another supplier in the following month. The payoff of a client matched with a good supplier is as before:

$$V^M = \frac{\alpha}{1-\beta}$$

The expected payoff of a client when searching is:

$$V^S = -c + (1-\theta)\beta V^S + \theta\beta V^M$$

Solving for V^S, we obtain:

$$V^S = \frac{\theta\beta\alpha - c(1-\beta)}{(1-\beta)(1-\beta+\theta\beta)} \tag{2.10}$$

Now we can ask ourselves the question: would A cheat a good supplier? If A cheats, he gets an instantaneous payoff of 1 as before but the continuation payoff is different from the first example: he now gets V^S. The non-cheating constraint now is:

$$1 + \alpha + \beta V^S \leq \alpha + \beta V^M = V^M \text{ or, written as in 2.9} \tag{2.11}$$

$$\pi(f, \tau, \varepsilon) = 1 \leq \beta(V^M - V^S) = EV(\tau, \varepsilon^*) \tag{2.12}$$

In this case, the value of the relationship $EV(\tau, \varepsilon^*)$ is the difference $\beta V^M - \beta V^S$. This is because a cheater still has a chance of forming a new relationship, but must incur a cost of search to do so.

Plugging equation 2.10 into the non-cheating constraint 2.11, after some algebra we can rewrite the non-cheating constraint as:

$$1 \leq \beta \frac{c+\alpha}{1-\beta+\beta\theta} = EV(\tau, \varepsilon^*) \tag{2.13}$$

We see that, if $\theta = 0$ (no replacement supplier) and $c = 0$, condition 2.12 boils down to 2.9. Consider for a moment what happens if $c = 0$. We see that $EV(\tau, \varepsilon^*)$ is a decreasing function of θ: the higher θ is, the lower $EV(\tau, \varepsilon^*)$ is. At the limit, if $\theta = 1$, condition 2.12 becomes:

$$1 \leq \beta \frac{\alpha}{1-\beta+\beta} = \beta\alpha \tag{2.14}$$

If the profit margin is less than 100%, the value of the relationship is too small and condition 2.14 is violated. There is still a penalty, however, because matching is not immediate – i.e., the cheater loses α for one period. This is what condition 2.14 says.

Now imagine that the cheating client can immediately find a new supplier and does not have to wait. In our example, this case can be represented by letting $\theta = 1$ as before and by setting $c = -\alpha$, meaning that the client makes a profit of α instead of incurring a cost of c. The non-cheating constraint becomes:

$$1 \leq \beta \frac{-\alpha+\alpha}{1-\beta+\beta} = 0$$

In this case, the relationship has no value. This is intuitively correct since the client can immediately replace the cheated supplier. Consequently, the fear of losing a supplier has value 0: the contract cannot be enforced by the fear of breaking a relationship.

2.4. Trust and Reputation

2.4.1. The different types of reputation

So far we have discussed various contract enforcement mechanisms. We showed that courts are not the only way contracts can be enforced. The fear of losing a valuable relationship can also induce economic agents to respect their contractual obligations. We new examine the role played by reputation.

The word reputation has been used with different meanings in the literature, so it is imperative to first distinguish between these meanings and agree among ourselves on a particular meaning. Reputation has first been used to describe relationships, as in the phrase 'If I cheat Jack, I will lose my reputation with him'. In this case, the word reputation simply refers to the relationship itself, which we have already studied. This is not the meaning of the word we will be using here.

A second meaning of the word refers to the type of the agent or good they produce, as when we say 'Toyota is a car manufacturer with a good reputation'. Here, presumably what we mean is that this car manufacturer produces reliable cars. We rely on the manufacturer's reputation to assess a hidden characteristic of the product we buy, namely how long it will operate without having to go to the garage for repair.

A third meaning of the word refers to the past behavior of an agent, as when we say 'Andersen lost its reputation when it helped Enron circumvent regulation on public securities'. In this case we are not talking about the quality of Andersen's service, but about the fact that they cheated in the past.

Today we will discuss the last two meanings of reputation and see how they interact to assist or undermine markets. Note that, unlike the first meaning, the last two meanings typically imply the sharing of information. The reputation of Toyota cars for reliability is based the experience of millions of previous buyers. By sharing their driving experience with us, they help us draw some inference about a hidden attribute of Toyota cars relative to other cars. Similarly, Andersen's cheating behavior is known to us because the newspapers have circulated the information to the public. It is because this information has been circulated that Andersen has lost other customers. If the information had been kept secret, it would only have affected Andersen's relationship with Enron and the Securities Commission.

Information sharing is the key distinction between the first and the third meaning of reputation: without information sharing, the third meaning boils down to the first. In contrast, the second meaning of reputation need not be based on information sharing. There are circumstances in which we experiment with various suppliers of goods and services to uncover their type. Think of checking out a new restaurant, for instance: we may do this on our own, without sharing the information with anyone. But this information, once acquired, can be shared if useful.

Sometimes, it is not in the interest of the buyer to share information with other potential buyers for fear of competition. Think of an employer-employee relationship, for instance. Employers typically 'try out' new employees to see whether they fit their new job. When they discover which employees are good and which are not, it is not in their interest to tell other employers because they may hire workers away from them. This is also true for banks and borrowers: it is not in a bank's interest to circulate information about who the good borrowers are.

What we now need is a framework with which we can combine and compare these concepts to discover what the role of reputation is in enforcing contracts and sustaining markets.

2.4.2. Sharing information about types

To illustrate the role of reputation, we expand on our earlier model to allow information sharing. Let us begin by assuming that clients share information about suppliers' types. Does it affect incentives to cheat? The answer is yes because, when clients share information on good and bad suppliers, they no longer have to incur the screening cost c.[13] The non-cheating constraint thus becomes:

$$1 \leq \beta \frac{\alpha}{1-\beta+\beta\theta} < \beta \frac{c+\alpha}{1-\beta+\beta\theta} \text{ if } c > 0 \tag{2.15}$$

We see that sharing information about types reduces incentives to respect contractual obligations. This is because it reduces screening costs and thus reduces the penalty for cheating. In this case, reputation in the second sense makes contract enforcement more difficult.

2.4.3. Sharing information about behavior

If we assume instead that suppliers share information about the past behavior of clients, we get the opposite results: reputation can make contract enforcement easier. To see why, suppose that (good) suppliers agree never to sell to clients who have not paid in the past. A cheating client therefore can never buy from a good supplier ever again. With this assumption, we are back to our first model, even though there are many agents. The short-term gain from cheating is, as before, 1. The long term loss from cheating is all future trade, i.e., $\sum_{t=1}^{\infty} \beta^t \alpha = \frac{\beta\alpha}{1-\beta}$. Consequently, in this case, we have:

$$EV(\tau, \varepsilon) + EW(\tau, \varepsilon) = \frac{\beta\alpha}{1-\beta} > \beta \frac{\alpha}{1-\beta+\beta\theta} \tag{2.16}$$

from which we see that sharing information about behavior raises the penalty for cheating and thus provides better contract enforcement incentives. Of course, as before, agents have to be sufficiently patient – β close enough to 1 – and the profit margin must be strictly positive and large enough. With this kind of contract enforcement, all gains from trade cannot be competed out: clients must make a positive margin so that it is in their interest to preserve their reputation, otherwise they will cheat.

The split of $\frac{\beta\alpha}{1-\beta}$ into its two components $EV(\tau, \varepsilon)$ and $EW(\tau, \varepsilon)$ is a bit arbitrary in this case. But it is useful to think of it as having two distinct parts: what the client would economize by not having to look for another supplier, which is given by $\beta\frac{c+\alpha}{1-\beta+\beta\theta}$, and the lost future trade opportunities because other suppliers would refuse to sell. We also have to take into account the fact that, since no supplier would agree to sell, it is in fact not in the client's interest to incur the screening cost c. We therefore obtain:

$$EV(\tau, \varepsilon) = \beta \frac{\alpha}{1-\beta+\beta\theta} \tag{2.17}$$

$$\begin{aligned} EW(\tau, \varepsilon) &= \frac{\beta\alpha}{1-\beta} - \beta \frac{\alpha}{1-\beta+\beta\theta} \\ &= \frac{\beta^2\theta\alpha}{(1-\beta)(1-\beta+\beta\theta)} \end{aligned} \tag{2.18}$$

To better see what is going on, it is useful to consider the case where there are no bad suppliers

[13] Strictly speaking, screening costs have to be incurred once – the first time a supplier is approached by a client. But in the long run, this cost is a vanishingly small proportion of expected average payoffs and can be ignored.

and $\theta = 1$. In this case, the value of the relationship and or reputation become:

$$\begin{aligned} EV(\tau, \varepsilon) &= \beta\alpha \\ EW(\tau, \varepsilon) &= \beta\frac{\beta\alpha}{(1-\beta)} \end{aligned}$$

This says that, in this case, the value of the relationship simply is the loss of profit margin in the next period, discounted to the present with discount factor β. This is indeed the value of the relationship that we have earlier derived for this case – see inequality 2.14. The value of reputation is the loss of all future trades from the next period onwards. Indeed, if there had been no information sharing, the client would have begun trading right away, missing only one trading period.

When $\theta < 1$, things are a bit more complicated because the client need not find a good supplier right away, in which case the value of the relationship is higher. This is what is captured in equations 2.17 and 2.18 above.

In the literature, the situation in which economic agents collude to exclude cheaters from future trade goes by various names. Kandori (1992) calls it a reputational mechanism or equilibrium. Greif (1993) calls it a multilateral punishment strategy. Sometimes it is also called collective punishment or exclusion.

2.4.4. Reputation and meta-punishment

Reputational punishment have received an inordinate amount of attention in the literature, so much so that it is customarily believed that collective punishments are easy to sustain and are pervasive in practice. Any evidence that economic agents share information is usually taken to imply that they collude to exclude cheaters, often without acknowledging the possibility that they exchange information about types, not about cheaters. In my own empirical work, I have found only limited evidence of reputational punishment. My interpretation for these findings is that a coordinated punishment strategy is difficult to sustain.

To illustrate this difficulty, let us consider the incentives faced by suppliers. Consider the following situation: one client A has cheated on – i.e., not paid – one of the suppliers. This supplier tells the others. Suppliers have agreed not to deal with cheaters. Clients know this. Now A approaches supplier B, claims it was all a big mistake, that he will not cheat anymore. The question is: is it in B's interest to refuse to deal with A?

To answer this question, let us imagine that B agrees to trade with A and let us consider A's incentive to cheat B. Since A is already blacklisted by all other suppliers, if he were to cheat on B he would not find any other supplier to trade with afterwards. Consequently, A's incentive to pay is the value of the relationship in a world where only relationships matter:

$$1 \leq \frac{\beta\alpha}{1-\beta}$$

But, as we have just seen, this is the same non-cheating constraint than the one faced by other clients. Consequently, A has as much incentive to pay B as any other client. From B's perspective, if the above constraint ensures that other clients would pay, then it also ensures that A will pay; A's future behavior is as reliable as any other client, irrespective of his past behavior.

Now consider B's incentive to trade with A. The alternative is to refuse to trade and wait for another client. If suppliers make zero margin, then B is indifferent. But if B makes a positive trade margin, then it is costly for B to refuse to trade and wait. Consequently, it is in B's interest to trade with A.

This problem is known as the meta-punishment problem: to incite agents to jointly punish cheaters, suppliers who refuse to punish must themselves be punished. If other suppliers could impose some social sanction on B, they might be able to force B to refuse and wait. The problem with meta-punishment in this case, however, is that trade between A and B need not be observable to other suppliers. Obviously, it would not be in the interest of either A or B to tell others that they are circumventing the sanction. Consequently, for meta-punishment to be implementable, suppliers must be able to observe each others' dealings. This requirement to some extent runs contrary to the requirements of competition, which assume that agents can undercut each other, etc. Circumstances under which meta-punishment is implementable for commercial contracts are thus difficult to satisfy.

In this context, if meta-punishment is impossible (either because social sanctions cannot be imposed or because dealing with cheaters is not observable), then collective punishment unravels. This means that, even though there is information sharing about past cheating, suppliers cannot coordinate their action to permanently exclude cheaters. We fall back on the N-agent case with non-cheating constraint:

$$1 \leq \beta \frac{c + \alpha}{1 - \beta + \beta\theta}$$

Greif (1993) offers a slightly different version of this model in which the supplier B finds it against his interest to trade with A. Without going through the details of his model, let me point out that Greif's model differs from ours in three important ways. First, suppliers are not matched at random with one client per period. Instead they pick a client of their choosing among all available clients (i.e., those not in a supplier-client relationship). This eliminates any loss from waiting since there is no waiting.

Secondly, Greif assumes that, in each period, a client-supplier relationship has an exogenous probability of ending. This assumption can be motivated as capturing changes in trade opportunities, technology, location, etc. This assumption is critical because it creates a wedge between the future payoff of a past cheater who, in the case of a single supplier deviating from collective punishment, would lose all future trade if the relationship with the deviating supplier B were severed; and the future payoff of non-cheaters, who would find another supplier if the relationship with B were severed. This means that the incentive contraint for non-cheaters has an extra reward for not cheating B.

Finally, Greif assumes that the seller/supplier extracts all the gains from trade. In other words, B gives to A the smallest possible α that ensures compliance. Since non-cheaters have an extra reward for compliance, they can be offered a smaller α than past cheater A and still be deterred from cheating. Hence the supplier B will naturally approach a past non-cheater when selecting a new client. In this case, the punishment of cheaters is self-enforcing without the need for a meta-punishment.

2.4.5. Self-enforcing collective punishment

It is also possible to make collective punishment self-enforcing in our model by introducing multiple types of clients. The difficulty of enforcing collective punishment originates in our assumption that all clients are identical. Consequently, the fact that a client has cheated in the past does not reveal anything about this client. It only says this client behaved strangely in the past. But his future payoff remains unchanged, and thus his incentive to cheat again in the future remains the same. If the threat of exclusion can deter cheating from all clients, then it can also deter future cheating by past cheaters. This is what creates an incentive problem for suppliers.

Things are different if there are several types τ of clients. Suppose there are two types of clients, good and bad. Good clients are as before. Bad clients are very impatient – low β – so that they cannot resist the temptation to cheat. Alternatively, we can assume that they are incompetent so that their profit margin is low or negative – low α. Consequently, bad clients have less incentive to pay. Whatever the case, they always cheat. We also assume that good clients are such that they have sufficient incentives to pay.

Now suppose that supplier B sells to a client C who does not pay. The behavior of C now serves as a signal regarding C's type: C is a bad client, a cheater. Now suppose that B tells other suppliers. Will other suppliers refuse to sell to C? The answer is yes because in this case past behavior predicts future behavior through inference about types. Consequently, if cheating is interpreted as a signal of a bad type, sharing information about past behavior will result in collective punishment even without meta-punishment. In this case, collective punishment is self-enforcing.

In the framework of our earlier model, this is equivalent to saying that the probability of payment depends on the client's type. Let the proportion of bad types in the economy be μ. If the type is unknown, the supplier must take a chance and her payoff is:

$$
\begin{aligned}
E(\Pi(f)|\Omega) &= \Pi(f)\Pr(payment) && (2.19)\\
&= \Pi(f)(1-\mu) + 0\mu && (2.20)
\end{aligned}
$$

However, if the type of the client is known, the probability of payment is either 1 (if the client is good) or 0 (if the client is bad).

This model can be extended to the case in which clients' type changes over time. This corresponds to a situation in which agents go bankrupt, or decide to retire from the business, and hence become very impatient. In this case, these agents traded reliably in the past and were thus regarded as good types, but subsequently became bad types and cheated. The same mechanism applies: if information about their behavior circulates, they will be excluded from future trade in a decentralized, self-enforcing manner.

The model could also be generalized to a situation in which the enforcement mechanism is guilt, not relationships or reputation. Certain agents are 'honest' in the sense that they would feel very bad if they cheated while others are 'dishonest' in the sense that they would not care. Past behavior can then be used to infer someone's honesty, i.e., innate or acquired capacity to self-inflict punishment by feeling guilty. These issues are discussed in detail in Platteau's 2 JDS papers published in 1996. This interpretation seems to be the most natural one, the one we would probably volunteer if asked to explain how we interpret cheating.

A related concept is that of shame, that is, the capacity to feel bad if discovered cheating. This is distinct from guilt in that guilt does not require public knowledge. Shame is a different feeling, one that can be influenced by information sharing about past behavior. Shame can serve as punishment as well. Note that, for suppliers to incur the cost of sharing information to shame people, they must derive some kind of morbid satisfaction from shaming others...

2.4.6. Incentives to share information about cheaters

We seem to have solved the problem of meta-punishment in a way that is intuitively appealing: after all, if someone has cheated, we intuitively think that there is something 'wrong' with them, i.e., we intuitively make an inference about their type. Does this imply that collective punishment is after all easy to obtain? Not so fast.

There are several other problems associated with collective punishment. First, it must be possible to identify agents unambiguously. It would be of not use to know that C has cheated if there is not way of telling who C is. One of the purposes of business registration (with identification of the owners and their address) is precisely to enable agents to be identified unambiguously. Informal sector firms which are nearly always unregistered usually complain they are not trusted. In my view, it is because it is much too easy for them to cheat and disappear.

An immediate corollary of the above is that, if precise identification is problematic, agents may use alternative methods to circulate information. One such method is the personal recommendation letter or personal introduction. Another is the creation of networks or communities in which individual agents are identified in person.

A second problem associated with information sharing is that cheated suppliers do not have an incentive to share the information with others. One reason is that, by telling other suppliers who the cheater is, they actually help their competitors. Even if this is not a consideration, sharing information requires an effort on the part of the cheated supplier without any direct counterpart. As a result it is difficult to incite suppliers to circulate accurate and current information about cheaters. Finally, agents may seek to 'capture' clients by telling other suppliers that they are cheaters.

In my empirical work, I have encountered examples of all of the above phenomena, so much so that I am convinced they explain why information sharing is not more prevalent.

2.5. Flexibility

Before moving on to networks, it is useful to say a few words about contractual flexibility. Even though flexibility by itself has nothing to do with information sharing, it clarifies the circumstances under which a cheater will be penalized collectively by all suppliers.

Until now, we have regarded breach of contract as a yes or no affair: either the client pays or he does not pay. In practice, there is a length of time during which it is unclear whether the client will pay. Also, in our interpretation of cheating as a signal about type, we have assumed that someone who does not pay is necessarily a bad type. In real life, things are not so clear because agents are faced by shocks ε.

Perhaps a slightly more accurate representation of reality would be to assume that bad clients cheat all the time but that good clients fail to pay with probability $\gamma < 1$ because of shocks. In this context, observing non-payment raises the probability that the client is a bad type, but does not make it entirely certain. To make this clear, consider the following situation. Supplier B has been selling to A for some time. It is common knowledge that all good clients have a probability of defaulting $\gamma < 1$ while, for bad clients, the probability is 1. Furthermore, each period, good clients have a probability σ of becoming bad.

Now suppose that B observes non-payment. What are the odds that A has become bad? It the probability that A has become bad over the probability that A cheats. The probability that A cheats is the sum of the probability that A becomes bad plus the probability that A does not become bad but faces a bad shock. The odds are thus:

$$\Pr(A \text{ is bad}|A \text{ cheated}) = \frac{\Pr(A \text{ is bad})}{\Pr(A \text{ cheated})} = \frac{\sigma}{\sigma + (1-\sigma)\gamma}$$

Suppose further that suppliers make a positive profit margin on each sale and let V^T denote the supplier's expected utility from selling to a good client.[14] For simplicity, we assume that it takes

[14] This is itself a combination of the probability of being paid, etc. But we do not need these details here.

exactly 1 periods for the supplier to find a good new client. The supplier must choose between keeping the client and risk losing 1 a second time, or reject the client and lose potential sales for 1 period.

Formally, after one case of cheating, it is in the interest of the supplier to continue selling to this client if:

$$\begin{aligned} \frac{\sigma}{\sigma+(1-\sigma)\gamma}(-1)+\frac{(1-\sigma)\gamma}{\sigma+(1-\sigma)\gamma}V^T &\geq 0 \text{ or} \\ (1-\sigma)\gamma V^T &\geq \sigma \end{aligned}$$

which, for a large enough V^T and a small enough σ, may be satisfied. In this case, it is in supplier B's interest to be flexible, that is, to allow A to skip payment once.

Flexibility does not last forever, however. Suppose that A cheats a second time immediately afterwards. What is now the probability that A is bad? Let E be 'cheat twice in a row' and let T_1 be 'A is bad' and T_2 be 'A is good'. Bayes law says that:

$$\Pr(T_1|E) = \frac{\Pr(E|T_1)\Pr(T_1)}{\sum_{j=1}^{2}\Pr(E|T_j)\Pr(T_j)}$$

We have:[15]

$$\begin{aligned} \Pr(A \text{ is bad}|\text{cheat twice}) &= \frac{1\sigma}{1\sigma+\gamma\gamma(1-\sigma)} \\ &= \frac{\sigma}{\sigma+(1-\sigma)\gamma^2} \end{aligned}$$

Since $\gamma < 1$ by assumption, we see that:

$$\Pr(A \text{ is bad}|\text{cheat once}) = \frac{\sigma}{\sigma+(1-\sigma)\gamma} < \frac{\sigma}{\sigma+(1-\sigma)\gamma^2} = \Pr(A \text{ is bad}|\text{cheat twice})$$

This is because a good guy is unlikely to cheat twice in a row. The supplier's incentive to continue selling to A now becomes:

$$\begin{aligned} \frac{\sigma}{\sigma+(1-\sigma)\gamma^2}(-1)+\frac{(1-\sigma)\gamma^2}{\sigma+(1-\sigma)\gamma^2}V^T &\geq 0 \text{ or} \\ (1-\sigma)\gamma^2 V^T &\geq \sigma \end{aligned}$$

Since the left-hand side is smaller, the supplier is less likely to sell again to A.

This argument can be extended to N cheating periods simply by raising γ to the Nth power. We have:

$$\lim_{N\to\infty} \Pr(A \text{ is bad}|\text{cheat } N \text{ times in a row}) = 1$$

with very rapid convergence to 1 if γ is small.

This demonstrate that suppliers may be flexible for a while but, after some time, will gradually begin to suspect that the cheater is actually a bad client and should no longer be traded with. My work on manufacturing firms in Africa demonstrate that firms often adopt this attitude.

[15] I have simplified this a bit. To be completely correct, we would have to allow for the fact that A cheated the first time because of a shock and the second time because he switched type. This would only complicate the math without changing the qualitative conclusion.

2.6. Discrimination and Networks

2.6.1. The building of trust

So far we have discussed information sharing and its effect on contract enforcement. We considered the possibility that clients and suppliers may come in various types. The existence of multiple types was found to be essential in most cases. In the absence of collective punishment, the presence of bad suppliers was critical to penalize cheating clients: it was the search cost and screening cost of identify another good supplier that induced clients to behave. Multiple client types were also found to be essential to support collective exclusion. We now revisit some of the issues raised by the presence of multiple types.

We first examine how trust can be built over time when there are many types of clients. Next we allow for observable characteristics that are correlated with hidden type and we introduce the concept of statistical discrimination. Finally, we discuss a variety of network effects that arise as a result of multiple types and information sharing within networks.

So far we have assumed that economic agents can ascertain each other's type by incurring a screening cost c. But we have not discussed where this cost comes from. To this we now turn.

2.6.2. Two types

Let us start with a simple example. Suppliers share information perfectly and collective punishment is enforced via meta-punishment. There are two types of clients, good and bad. Good clients have discount factor β_h while bad clients have discount factor β_l with

$$\beta_h > \beta_l$$

Being more patient, good clients value long-term gains more and thus have less incentive to cheat. To make this clear, we assume that:

$$\frac{\beta_l \alpha}{1-\beta_l} < 1 < \frac{\beta_h \alpha}{1-\beta_h} \tag{2.21}$$

This means that bad clients cannot be deterred from cheating (on a transaction of size 1 – see below) . Before trading (e.g., granting trade credit, accepting a check), suppliers wish to assess the client's type. The question now is how can suppliers force bad types to reveal themselves?

One possibility is for suppliers to propose a smaller transaction (e.g., grant a smaller amount of trade credit). By reducing the size of the transaction, the supplier reduces her exposure in case of non-payment. Let κ be the size of the transaction. What is the smallest transaction required to induce the bad agent to cheat?

The difficulty comes from the fact that if κ is chosen too small, bad clients may choose to pay the small transaction only to cheat better later. To see this, suppose that the supplier first sells κ and, if paid, sells 1. What is the bad client's incentive to mimic good behavior only to cheat better later? By cheating right away, the bad client gets $\kappa(1+\alpha)$. By cheating later, the bad client gets $\alpha\kappa$ now and $\beta_l(1+\alpha)$ next period. In order to induce the bad client to reveal his type right away, κ must be such that cheating now is better for the bad client:

$$\begin{aligned} \kappa(1+\alpha) &\geq \alpha\kappa + \beta_l(1+\alpha) \\ \kappa_1^* &\geq \beta_l(1+\alpha) \end{aligned} \tag{2.22}$$

It is easy to verify that $\kappa_1^* < 1$ since, by inequality 2.21, a size 1 transaction induces immediate cheating by bad clients.[16] It is also straightforward to verify that good clients would repay any size transaction since:

$$1 < \frac{\beta_h \alpha}{1-\beta_h} \Rightarrow \kappa < \frac{\beta_h \alpha \kappa}{1-\beta_h} \text{ for any } \kappa$$

Having found κ_1^*, we now easily derive c as the expected loss associated with the 'test' transaction κ^*. Suppliers have a unit profit margin of λ on repaid transactions. If the proportion of bad clients in the economy is σ, we have:

$$\begin{aligned} -c &= -\sigma\kappa_1^* + (1-\sigma)\lambda\kappa_1^* \\ c &= \kappa_1^*(\sigma - (1-\sigma)\lambda) \end{aligned}$$

A longer trial period may be better than a one-period trial period. Intuitively, insisting on a longer trial period enables the supplier to reduce κ_1^* because the impatient bad client has to wait longer before being able to fully cheat the supplier. For instance, if the trial period lasted 2 periods, the optimal level of κ would be given by:

$$\begin{aligned} \kappa(1+\alpha) &\geq \alpha\kappa + \beta_l \alpha\kappa + \beta_l^2(1+\alpha) \\ \kappa_2^* &\geq \frac{\beta_l^2(1+\alpha)}{1-\beta_l\alpha} \end{aligned}$$

which is clearly smaller than 2.22 since we can verify that:

$$\begin{aligned} \kappa_2^* &= \frac{\beta_l^2(1+\alpha)}{1-\beta_l\alpha} < \beta_l(1+\alpha) = \kappa_1^* \\ &\Leftrightarrow \frac{\beta_l}{1-\beta_l\alpha} < 1 \\ &\Leftrightarrow \frac{\beta_l\alpha}{1-\beta_l} < 1 \end{aligned}$$

which holds by assumption. When choosing how long the trial period should be, the supplier must trade lower exposure with reduction in trade volume in case the client turns out to be a good one. We leave this as an exercise.[17] In practice, trial periods in commercial relationships last from 3 to 6 months in sub-Saharan Africa. Trial periods of a similar length are found in labor contracts.

2.6.3. N types

Assuming only two types of client is obviously an oversimplification. Suppose instead that there are N types ranked $\beta_1 < \beta_2 < ... < \beta_N$. Let's partition these types into all the good types G

[16] To see this, note that 2.21 can be rewritten:

$$1 > \beta_l + \beta_l\alpha$$

If $\kappa = 1$, 2.22 becomes:

$$1 > \beta_l(1+\alpha)$$

which is clearly always satisfied if 2.21 is satisfied.

[17] As pointed out by a former student, Michael Best, a two period trial period is not time consistent/subgame perfect for the seller. After one period, the bad type reveals his/her type by cheating. Realizing this, the seller should switch to $\kappa = 1$ in the second period. Anticipating this, the low type buyer may not be deterred from cheating in the first period. The two-period trial is a Nash equilibrium but may be difficult to sustain in practice unless the supplier can credibly commit to a trial period of a given duration – e.g., a trial employment contract.

for whom

$$1 \leq \frac{\beta_h \alpha}{1-\beta_h} \text{ for all } h > t$$

and all the bad types for whom

$$\frac{\beta_l \alpha}{1-\beta_l} < 1 \text{ for all } h \leq t$$

The index t denotes the most patient bad type. Obviously, we know from 2.22 that type t is the hardest to force to reveal his type because it requires the highest κ^* to induce cheating.

One possibility would be for the supplier to set

$$\kappa^* = \frac{\beta_t}{1-\alpha}$$

in which case all bad types would reveal themselves in period 1, and the supplier would be fully informed. But this is an expensive strategy for the supplier because it requires a high exposure κ^*. If there are lots of cheaters in the economy, this is probably too expensive. Another approach would be to start small, first forcing the really bad types to reveal themselves. Then in the second period, the supplier would raise κ^* slightly to induce another batch of bad types to reveal themselves. And so on until all bad types have been revealed. It would also be conceivable to lengthen the trial period for each group in order to further reduce exposure, albeit at the cost of delaying full trade.

It should be intuitively clear that the exact shape of the screening strategy will depend on the precise distribution of types – i.e., their level of impatience and the proportion of clients of various types – as well as on the supplier's margin λ. We need not explore this further. All we need to remember is that this gradual screening process resembles what people usually call the 'building of trust'.

In practice, human beings are much better at inferring type than the crude model used here. They may rely on facial expression, verbalization, and various interpersonal interactions. This may help shorten the trial period and reduce exposure. These issues are areas of active research.

2.6.4. Types and observable characteristics

As we have seen, building trust is a lengthy and complicated process. To economize on it, economic agents may seek to use all the information at their disposal to speed up inference. Even though agents cannot observe someone else's type, they typically observe various external features that are correlated with the unobserved type.

For instance, a client may assume that a large firm is more reliable because its size indicates that other people have trusted it earlier. Of course, we know that this is not always true. Firms and individuals can organize sophisticated Ponzi or pyramid schemes by which they borrow increasing amounts to pay their mounting debt. The most infamous such scheme in recent history is the Enron scandal. Pyramid schemes led to civil war in Albania in the mid-1990s. Such schemes have also been documented in Nigeria and elsewhere.

Economic agents may also use easily observable personal characteristics such as gender, race, ethnicity, religion, appearance, age, and marital status to draw inference about their hidden type. In the literature this is called statistical discrimination. It is distinct from other forms of discrimination in that it is individually rational.[18] As a result, we expect statistical

[18] Other forms of discrimination such as the desire to only hire people from one's race are costly for the person doing it, because mutually beneficial opportunities for trade are ignored.

discrimination to be pervasive in all types of economic transactions where hidden type is at issue – labor, credit, insurance including health insurance, mortgages, house rentals, etc.

Statistical discrimination can take a variety of forms. The simplest form is simply to refuse to transact. We call this 'discriminatory rationing'. To illustrate this possibility, assume that there are two types of clients, good and bad. There are also two ethnic groups, B and W. The share of B clients in the population of clients is b. Without loss of generality, we assume that the proportion of bad types in the B population is higher than in the W population:

$$\sigma_B > \sigma_W$$

Now consider the decision of a supplier faced with a B client. The supplier can choose either to screen the client or to reject the client off hand and wait until next period in the hope of being faced with a W client. For simplicity, we assume that the screening exposure κ is the same for both populations of client and that screening is done in one period.[19] If the supplier screens the B client his payoff is:

$$V^B = \sigma_B(-\kappa + b\beta V^B + (1-b)\beta V^W) + (1-\sigma_B)(\lambda\kappa + \beta\frac{\lambda}{1-\beta}) \tag{2.23}$$

where λ is the supplier's profit margin and β his discount factor. The first term is what happens if the client is bad, in which case the supplier loses κ and is matched with a new client in the following period, who could either be B or W. The second term is what happens if the client is good, in which case the supplier makes his margin on the κ transaction today and gets λ thereafter. A similar equation can be written for V^W:

$$V^W = \sigma_W(-\kappa + b\beta V^B + (1-b)\beta V^W) + (1-\sigma_W)(\lambda\kappa + \beta\frac{\lambda}{1-\beta}) \tag{2.24}$$

It is immediately obvious that $V^B < V^W$ since $\sigma_B > \sigma_W$.

Now if the supplier refuses to screen B clients, the payoffs are:

$$\bar{V}^B = b\beta\bar{V}^B + (1-b)\beta\bar{V}^W \tag{2.25}$$

$$\bar{V}^W = \sigma_W(-\kappa + b\beta\bar{V}^B + (1-b)\beta\bar{V}^W) + (1-\sigma_W)(\lambda\kappa + \beta\frac{\lambda}{1-\beta}) \tag{2.26}$$

The question we now ask is: is it ever in the supplier's interest to refuse screening B clients simply on the basis of their external characteristics, not on the basis of their type which is unobservable? This is an important question because such discrimination is bound to be deeply resented by good B types.

To answer this question, we first solve the system made of 2.23 and 2.24. This yields the value of V^B. We then solve the system made of 2.25 and 2.26. This yields $\bar{V}^B$. We then compare V^B to $\bar{V}^B$ and check whether it is ever the case that the second is larger than the first. Skipping the tedious algebra and assuming for simplicity that $\kappa = 1$, we obtain an expression of the form:

$$V^B - \bar{V}^B = [\lambda(1-\sigma_B) + \sigma_W(1-b)\beta - \sigma_B(1-b\beta)]\,M$$

where M is a long expression guaranteed to be positive. Hence, the sign of $V^B - \bar{V}^B$ depends upon the sign of the expression in brackets. It is clear that this expression increases with λ: the higher the supplier's margin, the more he has to lose by delaying the chance of finding a

[19] Because $\sigma_B > \sigma_W$, however, the supplier may opt for a different screening strategy for each group, e.g., choosing a longer screening period and lower exposure for the B clients.

good client. The expression is also unambiguously decreasing in σ_B: the higher the proportion of bad clients in the B population, the less likely the supplier is to give B clients a chance to prove themselves. More precisely, we see that what matters is the difference between σ_W and σ_B (since β is close to 1, $(1-b)\beta \simeq 1-b\beta$): the larger the gap between the two, the more likely the supplier is to refuse screening. Finally, note that σ_B does not have to be much larger than σ_W for discrimination to arise: if the supplier is very patient but the margin λ is small, even a very small difference between σ_B and σ_W will result in discrimination.

To see the effect of screening cost κ, it is worthwhile considering the complete formula for $\kappa \neq 1$:

$$V^B - \bar{V}^B = \{\kappa\left[\lambda(1-\sigma_B) + \sigma_W(1-b)\beta - \sigma_B(1-b\beta)\right] + (1-\kappa)\beta\lambda(1-\sigma_B)\}M$$

The expression now has two terms, the second of which is always positive. From this it follows that, when the cost of screening κ is small, the supplier always prefers to screen. The reason is that not screening delays the time at which a good client is found.

An immediate policy implication is that reducing the cost of identifying good B agents should eliminate discrimination. Example: credit reference agency, certification programs, recommendation system. Of course, by helping good B agents to distinguish themselves from bad B agents, it makes it harder for bad B agents to be screened. Example: credit programs to microenterprises.

It possible to further analyze this model to demonstrate that, for certain parameter values, price discrimination obtains instead of discriminatory rationing. In this set-up, B agents have to accept a lower profit margin α (and thus guarantee suppliers a higher profit margin λ) in order to convince suppliers to screen B agents. In yet another set-up, suppliers agree to screen only a proportion of B agents, but not all. These various cases are discussed in my paper on Ethnicity and Networks in African Trade.

2.6.5. Negative feedback and affirmative action

Discrimination is inequitable because it deprives good B clients from the same opportunities as good W agents. It can also affect the behavior of B agents. To see why, suppose that the type of agents depends on an investment C. This could be an investment in education, capacity, self-discipline, etc. It is optimal for an agent to invest C if $C \leq R$ where R is the return on the investment. For W agent guaranteed to be screened, we have:

$$C \leq R_W = \frac{\beta\alpha}{1-\beta}$$

Discrimination affects the return to the investment. In our extreme example of discriminatory rationing, B agents who have made the effort of acquiring the good type reap no reward since they are never screened and thus never get to earn the margin α. In their case, $R_B = 0$.

It type is genetically determined, discrimination has no effect on the distribution of types across the two populations. But if acquiring the good type is costly, if would be irrational for B agents to make the effort since they would get no return on their investment. If this is the case, discrimination becomes self-fulfilling: B agents are bad *because* suppliers believe they are bad. Discrimination becomes its own justification. Coate and Loury analyze this situation in detail with a model of discrimination in job assignment.

One possible solution to this problem is 'affirmative action', that is, policy that forces suppliers (in our case) to trade with a certain proportion of B clients. Defined in this way, affirmative action is a very pervasive policy instrument used by governments, donor agencies, and NGOs alike. Any kind of targeted program is de facto an affirmative action program as we have defined it here.

The question is whether targeting solves the problem. If types are exogenously – i.e., 'genetically' – determined, it does: forcing suppliers to sell to B clients will force them to screen B potential clients until they find good B types. But if types are the result of individual investment, perverse results can obtain. Suppose for instance that no B client acquires the good type but that suppliers still have to sell to B clients. Then B agents can enjoy a positive payoff by being offered a κ transaction and refraining to pay. This payoff is obtained without making any investment in acquiring the good type. If this payoff is large enough relative to the cost of acquiring the good type, all B may choose to remain bad.

This kind of situation can be called a patronizing behavior: affirmative action artificially raises the welfare of B agents without improving their quality. If the policy were removed, suppliers would immediately go back to the original discriminatory rationing.

A virtuous circle may also arise if the benefit from κ are small enough relative to the gain of α forever. In this case, many B agents will find it optimal to invest in the good type. Coate and Loury discuss this issue in more detail.

2.7. Network effects

We have shown that observable differences among individuals can lead to discrimination and thus differential treatment among good types. Observing differential treatment does not, however, imply discrimination. There are other processes that can generate differential treatment. One of them is network effects.

To understand how this works, let us go back to our world with no observable characteristics. All potential clients are observationally equivalent. Let us assume instead that suppliers share information about clients' types. We ignore the various problems that arise in this context and assume that the information is passed on accurately. But we now assume that information does not circulate among all suppliers. Some suppliers, which we call the 'in' group, share information among themselves, while other suppliers, which we call the 'out' group, do not share information with anyone.

Suppliers that are 'out' essentially behave as earlier, so we ignore them for now. Focus instead on an 'in' supplier who is approached by a new potential client. Either the client is known to the 'in' group – we call this an 'in' client – or the client is not known – an 'out' client. If the client is an 'in' client, the supplier learns the client's type from his information sharing group. Either the 'in' client is good or bad. If the client is good, no need to screen: the supplier offers to trade immediately. If the client is bad, the client is rejected right from the start.

If the client is an 'out' client, the supplier has to screen to find out if he is good or bad. Screening costs κ. Alternatively, the supplier can reject the 'out' client and wait until the next period in the hope of being approached by a good 'in' client. Intuitively, if the screening cost κ is high and the chance of meeting a good 'out' client next time is high, waiting will be more advantageous than screening the 'out' client. Hence, information sharing can hurt good 'out' clients who do not get screened, as in the discriminatory rationing case.

To show this formally, we again consider the supplier's decision to screen or not to screen 'out' clients. We keep much of the same notation as in the previous sub-section: B stands for 'out' client; W stands for 'in' client. This time, the proportions of good clients are the same in both, i.e., $\sigma_B = \sigma_W = \sigma$. The proportion of 'in' clients in the economy is b. If the supplier screens 'out' clients we have:

$$V^B = \sigma(-\kappa + b\beta V^B + (1-b)\beta V^W) + (1-\sigma)(\lambda\kappa + \beta\frac{\lambda}{1-\beta}) \tag{2.27}$$

where as before λ is the supplier's profit margin and β his discount factor. The first term is what happens if the client is bad, in which case the supplier loses κ and is matched with a new client in the following period, who could either be B or W. The second term is what happens if the client is good, in which case the supplier makes his margin on the κ transaction today and gets λ thereafter. A similar equation can be written for V^W:

$$V^W = \sigma(b\beta V^B + (1-b)\beta V^W) + (1-\sigma)\frac{\lambda}{1-\beta} \tag{2.28}$$

In contrast to 2.24, we note that the supplier does not lose κ if the client is bad and earns the full λ immediately.

If the supplier refuses to screen B clients, the payoffs are:

$$\bar{V}^B = b\beta\bar{V}^B + (1-b)\beta\bar{V}^W \tag{2.29}$$

$$\bar{V}^W = \sigma(b\beta\bar{V}^B + (1-b)\beta\bar{V}^W) + (1-\sigma)\frac{\lambda}{1-\beta} \tag{2.30}$$

Again, we can solve each system of equation separately (they are just linear equations in V^B and V^W) and compute the difference between $V^B - \bar{V}^B$ which, for $\kappa = 1$, simplifies to:

$$V^B - \bar{V}^B = [\lambda(1-\sigma) - \sigma(1-\beta(b+(1-b)\sigma))]\, Q$$

where Q is a positive expression. We see that, the larger σ is, the less likely the supplier is to screen the 'out' client. It is also possible to verify that if $\kappa = 0$, it is always optimal to screen and that $V^B - \bar{V}^B$ decreases in κ.

The bottom line is that networks can generate discriminatory rationing. In my paper on Ethnicity and Trade, I show that networks can also generate discriminatory pricing.

2.8. Summary

From this overview of the contract enforcement problem, we have learned that penalties for breach of contract play a crucial role in making exchange possible whenever delivery and payment are not instantaneous. This applies to virtually all types of exchanges. Credit, insurance, and labor transactions are examples of transactions which inherently involve an element of time. This is true also of most services which imply some time element (e.g., haircut, cab ride, brokerage, tailor): the service provide fears providing the service without being paid, while the client fears paying without receiving the service.

Most transactions on goods and services also involve an element of time whenever payment is not instantaneous and the quality of the good cannot be inspected on the spot. For instance, when I purchase a tin of peas or a bottle of milk, I cannot assess the quality of the good inside the package. Consequently, for me to buy I must either trust the trader or trust the brand. This implies some kind of relationship either with the brand or the trader.

Breach also becomes possible if the client wishes to pay by check or credit card, because fraud is possible. This problem has become particularly acute in internet transactions where the percentage of fraudulent purchases is much higher than in regular transactions.

Various mechanisms can discourage breach of contract. Some have to do with guilt, some with legal or extra-legal penalties, some with the value of relationships. In this respect, we have seen that relationships have a value when there are no alternatives or when alternatives are costly to find.

Next, we discussed the role played by information sharing. We have seen that information sharing does not, by itself, improve contract enforcement, as is often believed. In fact, in some instances – sharing information about good suppliers – it may even weaken contract enforcement.

We showed that if cheaters are excluded from trade with all suppliers, enforcement of contracts is easier. We showed that the value of combined value of lost relationship and lost reputation is the same as the value of losing the relationship in a single supplier economy.

We also discussed the difficulty of organizing the collective exclusion of cheaters from future trade. One possibility is to rely on meta-punishment, that is, to punish suppliers who trade with clients who have previously cheated on other suppliers. But this is difficult to enforce. Another possibility is to interpret cheating as a signal about the client's type. In this case, exclusion from future trade is self-enforcing.

Finally we discussed what arises when clients face shocks. We showed that suppliers may choose to be flexible, that is, to allow a client not to pay for one period. But as clients repeatedly fail to pay, suppliers would revise their belief and eventually refuse to trade with the delinquent client.

Moving on to discrimination and network effects, we have seen that the presence of multiple types in the economy singularly complicates the operation of markets. The building of trust is like peeling an onion: it can be understood as a gradual discovery process in which temptation force various 'layers' of bad clients to reveal themselves. In the end, only the good clients are left. Crooks are the very patient bad agents who manage to gain other people's trust only to cheat them in a larger scale (e.g., pyramid schemes).

When types are correlated – even mildly – with observable characteristics, discrimination is likely to arise. Discrimination is likely to be particularly acute if screening costs are high and suppliers make low margins. Discrimination can be self-fulfilling if agents invest in their type. Affirmative action can in some circumstances get rid of the bad incentive to invest. But in other circumstances, it generates a patronizing behavior in which the disadvantaged population takes advantage of affirmative action.

Finally, we showed that networks can have effects similar to those of discrimination. Consequently, we should not necessarily interpret discriminatory rationing as evidence of statistical discrimination. We now discuss briefly policy implications for development.

2.9. Policy implications and market support institutions

I have argued that development is characterized by an expansion in the range of exchanges that become covered by the market. I have also argued that in any market exchange – except perhaps for the most rudimentary 'flea market' forms of exhange – there is room for breach of contract, that is, there a chance of being cheated. For market exchange to expand, people must trust each other. What provides this trust is market institutions. Here are some observations relevant for policy:

- For some people, saying that market institutions matter for development is synonymous to improving laws and courts: fund and equip courts; adopt 'modern' laws – sometimes advocate adop Common Law.

- Many market transactions are beyond the reach of court adjudication. Many transactions enforced on the basis of reputation and relationships. This may give the impression that laws, courts, and other formal institutions and organizations are not important. But market support institutions – of which laws and courts are the cornerstone – are very important to make reputation and relationships work. Here are some examples.

- Reputation mechanisms can be formalized/reinforced by modern market institutions:
  - Publicization of court cases (historically important; still basis for D&B). Computerize court registries; keep access fee low.
  - Credit reference agencies (eg Dunn and Bradstreet).
  - Credit rating agencies (eg Moody's).
  - Vetting agencies (eg university or training center that 'vet' their students).
  - Customer feedback (eg Ebay)
- Reputation assumes that identity can be established:
  - ID cards; Social Security Number or National Insurance Number.
  - Business registration; this requires a system of physical addresses (not just POBox).
  - Registrar of companies, or NGOs, with power/funding to investigate and verify.
  - Ebay account: easy to create; but an account with thousands of past customer feedback becomes valuable and worth protecting.
  - Phone number for Mpesa.
- Reputation assumes that identity is protected:
  - Identity theft: if too much, internet commerce will unravel.
  - Trademark and brand names (eg McCurry)
  - Impersonating someone else is made an offence and punished by law.
- Reputation helps established firms and agents, not new agents. Institutions can be put in place to favor the formation of new relationships:
  - Business associations, private and public; this assumes that these associations are not seen as threatening by the political power (eg Ghana under Rawlings, which hunts, corruption campaigns aimed at political rival base)
  - Affirmative action and targetting: to 'co-opt' economic agents that find it difficult to 'enter' and hence cannot develop a reputation (eg, credit for female enterpreneurs, affirmative action for minorities in recruitment and public contracts.
  - Targetted programs to help school leavers get their first job.
  - Creation of registries for those who currently do not have an 'identity' (eg, registry of informal sector enterprises).
  - Emancipating those who currently cannot be legally accountable (eg emancipation of women and their rights to run a business and borrow money) – only then can they develop a specific legal and economic identity.
  - Helping local enterpreneurs break into foreign markets; invite them on trade tours, etc.
- Formal institutions (public and private) play a paramount role in certain key markets:

- Protection of property rights:
  * theft
  * crooks
  * corruption
  * extorsion/racketeering
  * expropriation
  * land registry
  * registry for vehicles, machinery (sometimes private)
  * id numbers, tags, bugs, and implants
- International transactions:
  * Letter of credit, correspondent banks, and the international system of payment.
  * Transnational corporations.
  * SGS and remote inspection of goods.
  * Credit reference and credit rating agencies: to get your firms known abroad.
- Specialized markets for super-liquid transactions: here the potential is enormous
  * stock market
  * commodity exchange
  * art auctions
  * secondary market for mortgages
  * Regulation is essential to keep crooks out and to ensure participants are liable.
  * Standards must be set in place so that physical goods can be precisely described and sold without the need for physical inspection.
  * Standards must be assessed by independent auditors (eg, auction houses for art objects, captain of vessel for connaissement,
  * Warehouses receipts and bonded warehouses.
  * Internet dissemination of information.

- Collaterizeability:
  - All the assets of an individual can be seized to serve his/her debts. Hence all assets are collateral.
  - But lenders can seek additional protection:
  - eg: Tax administration, social security administration usually have seniority by law.
  - eg: workers usually have seniority by law for payment of unpaid wages.
  - eg: AFTER the above, lender can earmark certain assets to serve their debt in priority, eg mortgage, vehicle lease, garnishing of wages and idemnities (eg pension)
  - This is only possible if institutions have been put in place to make these goods/entitlements 'collaterizeable'
  - Given this, lender may worry that borrower would liquidate the good, hence the need for a registry (eg. of mortgages, of lien on vehicles and machinery, which also assume that identification is unambiguous through id tags or land survey)

- Bankruptcy: laws to protect those that have failed so that they can start afresh.

  - corporate bankruptcy and limited liability: a way to protect personal assets.
  - problem if also seek public funds: need for regulation (eg public accounts, external auditors)
  - personal bankruptcy?

- Etc.

  What this list shows is that (1) there are many market support institutions other than laws and courts; and that (2) even though laws and courts may not enforce many transactions directly, they are essential for the functioning of nearly all the formal market support institutions.

3. Lecture 4. Applications to micro-finance

References:

Zeldes (1989), "Consumption and Liquidity Constraints: An Empirical Investigation", *J. Polit. Econ.*, 97(2): 305-343

Carroll (1992), "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence", *Brookings Papers on Econ. Activity*, 2: 61-156

Deaton (1991), "Saving and Liquidity Constraints", *Econometrica*, 59(5): 1221-1248

Hall (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence", *J. Polit. Econ.*, 86(6): 461-81

Besley, Coate and Loury (1993), "The Economics of Rotating Savings and Credit Associations", *Amer. Econ. Rev.,* 83(4): 792-810

Besley and Coate (1995), "Group Lending, Repayment Incentives, and Social Collateral", *J. Devel. Econ.*, 46(1): 1-18

Ghatak (2000), "Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect", *Economic Journal*, 110(465): 601-31

Ashraf, Karlan and Yin (2006), "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines", *Quarterly Journal of Economics*, 121(2): 635-672

Ambec and Treich (2007), "Roscas as financial agreements to cope with self-control problems", *J. Devel. Econ.*, 82: 120-37.

3.1. Introduction

In this lecture I present examples of applications of the market exchange framework to micro-finance. Rather than focus on a small number of published papers, I present a personal overview of what I perceive are key concepts – not always well presented in the literature.

3.2. Poverty trap and the market for investable funds

Ever since the 1950 and the work of Nurkse and Hirshman, there has been an enormous interest in credit in the development literature. The basic insight is best captured by the phrase coined by Nurkse: 'the poor are poor because they are poor'. Poor people have insufficient funds to invest. If there are profitable investment opportunities open to them but they cannot invest, they will not be able to earn the return to these investment opportunities and they will remain poor. This is the simplest expression of the poverty trap hypothesis.

There are several implicit assumptions in this reasoning. First, it assumes implicitly that the profitable investment opportunities could not be captured by others. If they can, then poverty is a problem in terms of equity – e.g. someone else with money will make the investment – but not really in terms of efficiency or development – profitable investments will be made and the country will grow. This literature typically assumes investment opportunities that are individual specific – perhaps as a result of ownership of or control over valuable resources or because of individual specific skills and knowledge. These valuable resources can be underground oil or unpaid family labor.

Sometimes, it is implicitly assumed that the investment opportunity is country-specific but not individual specific. The early literature – i.e., immediately before and after decolonization – kind of regarded the two as interchangeable because developing countries were thought of as having no indigenous wealthy class. Furthermore, 1950's authors implicitly understood that a foreign investor risks the risk of being expropriated – as had just happened in communist countries – and thus may not have seen foreign investors as a suitable substitute for a local wealthy class. Things have changed since then and there are wealthy individuals in all countries these days, including Africa. In this case, it is not immediately clear why high return investment opportunities in a country would not be taken advantages of – except for very large investments (e.g. a pipeline). Perhaps the local wealthy class does not have the necessary skills and knowledge to undertake the investment. But then the problem is not lack of funds but lack of expertise. That's not what the poverty trap hypothesis is about.

Secondly, the simple poverty trap hypothesis implicit assumes that markets are unable to channel funds to investors. This is typically expressed in terms of 'credit market failure' although it could just as well be seen as a 'equity market failure'. In the first lecture I illustrated how in the absence of credit, saving and investment depend on the consumer's preferences – an example of failure of separability. Without credit market we gen an Euler equation, which has marginal utility in it – and hence consumption levels (i.e., poverty). A poor consumer values consumption today more than consumption tomorrow. As a result there is underinvestment. This was illustrated in lecture 1 using a two-period (human capital) investment model. With a perfect credit market, we get the standard result that the marginal rate of return to capital equals the interest rate – utility does not matter. For memory, the model was:

$$\max_{\{K_t, W_t\}} \sum_{t=0}^{\infty} \beta^t V(C_t) \text{ subject to}$$
$$C_t = Y(K_t) - rK_t - W_{t+1} + (1-r)W_t$$

where W_t is the net wealth of the household, and r is the interest rate. I have assumed that the household can save and earn an interest r. I have also omitted depreciation δ from the notation – think of it as included in $Y(K)$. Since only K_t enters the budget constraint, not K_{t+1}, the Euler equations for K simply is in each period:

$$\frac{\partial Y(K_t)}{\partial K_t} = r$$

The problem is that, without uncertainty, it is difficult to understand why there would not be a credit market in the first place: the lender can perfectly predicts the return to the investment, hence the only risk that exists is opportunistic breach of contract, i.e., the borrower can pay but chooses not to repay. This can – and probably should – be solved by improving market institutions. In general, since all of people's assets can be called upon to serve their debts, the future income from a project can be used as collateral. There isn't really a problem of lack of collateral. It is like somebody borrowing to buy a house: the house itself, that is, the product of the investment serves as collateral for the loan. Same observation for a car loan or a loan guaranteed by a lien on machinery. Thus the only way to make sense of credit market failure in this context is not to think of it as a problem of lack of collateral – by assumption, lenders know which borrowers have a good investment opportunity and hence future income/assets from the project are the collateral – but rather a problem of contract enforcement – borrowers can dissimulate or dissipate the funds before the lender can recoup them. The only problem in this case is dishonest borrowers.

In the past two lectures we have discussed how to provide deterrent for opportunistic breach of contract. If the cause of credit market failure is that poor borrowers cannot be forced to repay, it seems foolish to force lenders to lend to them. Yet this is what is often advocated – e.g., incite lenders to lend to people with insufficient wealth and income; this, combined with an active secondary market for mortgates, resulted in the sub-prime crisis and a major world economic crisis. Similar lending programs have been implemented in many developing countries, e.g., India forcing banks to lend to farmers in spite of the fact that it is de facto impossible to repossess land in rural India; as a result, many such loans are not repaid. Similar programs were implemented in Indonesia, etc.

What happens if we introduce uncertainty? Let the model remain as before:

$$\max_{\{K_t,W_t\}} \sum_{t=0}^{\infty} \beta^t V(C_t) \text{ subject to}$$
$$C_t = \widetilde{Y}(K_t) - rK_t - W_{t+1} + (1-r)W_t$$

where $\widetilde{Y}(K_t)$ is now a random variable. This is the kind of model studied by Deaton and Zeldes at the beginning of the 1990's. In general it is reasonable to assume that $\Pr(Y = 0) = \tau > 0$ – there is some probability that future income will be zero, for instance because of death or disability of the borrower. We also assume that $C_t \geq 0$ – people cannot 'regurgitate' consumption, that is, produce value by having negative consumption. If we make these two assumptions, then I claim that $W_t \geq 0$ at all t. This happens even though we have NOT assumed a missing credit market.[20] What is going on?

The reason for this result, which I have first seen in Carroll (1992), is that the above mathematical model does not allow for default. It requires that debts be paid in ALL states of the world. To understand Carroll's logic, consider a finitely lived borrower who lives until T. Imagine that this person borrowed B at time $T-1$. His wealth or net worth at $T-1$ is thus $W_{T-1} = -B$. This means that this borrower must repay B out of income $\widetilde{Y}(K_t)$. But we have assumed that $\Pr(Y = 0) = \tau > 0$. This means that with probability $\geq \tau$ the loan B will not be repaid. This violates the final period budget constraint. Hence it is not possible for this individual to have a negative net worth at time $T-1$.

Now consider time $T-2$. We have already established that the person cannot borrow at $T-1$, and hence cannot postpone repayment at $T-1$ by carrying over a negative net worth to period T. This means that any amount lent at $T-2$ must be paid at $T-1$. But then the same argument as above applies since we still have $\Pr(Y = 0) = \tau > 0$: net worth W_{T-2} also cannot be negative. And so on: recursive induction implies that, even though there is no explicit credit constraint of the form $W_t \geq 0$ in the model, nevertheless $\Pr(Y = 0) = \tau > 0$ implies $W_t \geq 0 \forall t$. This is an often overlooked result – even Deaton's pathbreaking work on precautionary savings does not mention it (and there is some suggestion that at the time he wrote his Econometrica paper on precautionary savings, he had not fully grasped it).[21]

What is the point of this discussion? The point is that, in a generic sense, credit markets can only exists if we allow for bankruptcy and default. Put differently, credit can only exists if lenders accept that, with some probability, loans will not be repaid. The problem is that standard models have nothing to say about repayment probabilities since they do not allow for default. To understand repayment probabilities, one has to understand incentives to repay. Hence my earlier emphasis on market institutions.

If we allow for default in certain states of the world, we enter a world of contingent debt contracts: loans are repaid in most states of the world but not in some. This means that the lender de factor shoulders part of the investment risk. If this is the case, why not opt for a fully contingent repayment, that is, for equity funding? This solves the problem of contingent debt by making repayment automatically contingent on the success of the investment. This is

[20] Of course it is possible to have borrowing with $W_t > 0$, e.g., borrowing to buy a house, in which case net worth remains positive. This is not what poverty traps are about so I do not discuss this case further.

[21] Hall's 1978 paper on certainty equivalence also fails to notice this problem; in his case the problem is that he does not impose the condition that $C_t \geq 0$.

the domain of 'angels' and venture capital. Here too there is an enforcement problem, however: the investor can dissimulate or dissipate equity returns. Unlike in loan contracts, contingencies have to be assessed – i.e., the profit of the firm must be measured accurately – in ALL states of the world. Hence the need for more oversight/monitoring. But monitoring is costly. With transactions costs, we would expect only the highest return projects to be funded by venture capital.

Loan contracts are simpler to implement in most states of the world since payment is, most of the time, not contingent. It only becomes contingent in case of default. One elegant solution in this case is to organize monitoring/oversight ONLY in cases of default. This is what bankruptcy is about: it allows courts to investigate defaulting firms ex post, and to charge managers with embezzlement or fraudulent bankruptcy if funds were indeed dissimulated or diverted. This approach economizes on monitoring and also automatically appoints an external monitor, the court. Nice solution.

It follows that if you want to understand why some people cannot borrow, study how default and bankruptcies are handled by formal and informal market institutions. For instance, some countries allow for personal bankruptcy (e.g., US), others not (e.g, France). You will find that typically poor borrowers can avoid penalties (except when they borrow from loan sharks, who will resort to physical threats and punishment) – and this is what makes them unattractive borrowers in the first place. The literature calls this being 'judgement-proof'.

One may be tempted to solve the judgement-proof problem by introducing stricter penalties for default – e.g., prison, as was the case in the UK last century. The problem with this approach is that, in a probabilistic sense, there will be people who, through no (or little) fault of their own, find themselves unable to pay. These people will go to jail nevertheless. We may not want that. Another solution is personal indenture: people who cannot repay become 'slaves' – they must work for someone until their debt is repaid. This solution was applied in Europe, e.g., in the 17th and 18th century. This system is still in use (although illegal) in some countries. The same comment applies.

3.3. ROSCAs

A ROSCA (or rotating savings and credit association) is a common contract found in many places of the world. Also called *tontine* or *chit fund*. In its simplest form, it is a contract by which N individuals agree to pay a fixed amount m over N fixed periods, called a rotation. In each of these N periods, all contributions Nm – the pot – are given to one member selected at random among those who have not yet received the pot during that rotation. At the end of a rotation, each member has received the pot once and only once. The process is then restarted. Members can enter or exit only at the end of a rotation.[22]

Roscas are often used by people who wish to save and have a given target of Nm, e.g., traders to replenish their working capital, housewives to purchase household durables, or civil servants to pay for the pilgrimage to Mecca. These individuals could in principle save on their own, putting m aside each period. By saving alone, if would take N periods to save Nm. As argued by Besley, Coate, and Loury (1993), in expectations term a rosca speeds up the time to investment.

[22] There are more complicated roscas, such as roscas in which members can bid for the pot by offering to leave some money in the pot in exchange for getting the pot earlier, i.e., by offering to take $Nm - b$, where b is given to the person whose turn it was (if the person has already been selected) or b is shared among those members who have not yet received the pot during the rotation. These roscas are called bidding roscas. I do not discuss them here.

For instance, say that $N = 10$. Assume for now that people save the same amount in autarky and with the rosca. By forming a rosca, it takes on average $E[1 + 2 + 3 + ... + 10]/10 = 5.5$ periods to save, which is less than 10. If people discount the future ($\delta > 0$) their utility is higher – irrespective of the form of the utility function:

$$e^{-10\delta}U(Nm) < e^{-5.5\delta}U(Nm)$$

for any Nm. Only if $\delta = 0$ are individuals indifferent.[23] Note that for this result it is essential that each person be guaranteed to receive the pot once in each rotation. Simply drawing at random each period, with replacement, would violate the above reasoning for someone sufficiently impatient/risk averse.

What is the best way to randomize the allocation of the pot? Is it better to randomize once and keep the same order, or is it better to re-randomize in each rotation (assuming that the people do not change)? It turns out that it is in general better to randomize once and to keep the same randomized order thereafter. To see why, imagine the contrary, i.e., the order is re-randomized in each rotation. Consider the first round of the rotation. Individual i is selected to win the pot. In round 2 this individual has to decide whether to stay in the group or not – i.e., whether to default or not. This person now has to wait on average $9 + 5.5 = 14.5$ periods before getting the next pot. On her own, this person would only have to wait 10 periods. There is therefore an incentive to default. If we keep the same order, however, individual i has to wait exactly 10 periods before getting the pot next time. It is therefore easier to deter default if we keep the same order across rotations (see Ambec for a discussion).

This, however, brings out another issue: why would people stay in a rosca if all the benefits are earned during the first rotation and there are not benefits in subsequent rotations? If there are (even small) transactions cost in participating to a rosca, we should observe people joining a rosca only once and then dropping out. Yet this is not what we normally observe: people often continue.

This observation also has implications for contract enforcement. Given the small sums typically involved, it is reasonable to assume no external/judicial enforcement. Reputation also has no bite in this case since after the first rotation people can do as well on their own. In a repeated arrangement of this kind, we would normally expect that some kind of self-enforcement, i.e., people respect their obligations because they expect future benefits. How then is the rosca enforced in the first place, i.e., why don't people simply leave immediately after receiving the pot? But then if they would leave immediately after receiving the pot because there are no incentives to continue, this should have been anticipated and roscas should not exist in equilibrium. Yet they do.

Does the BCL model solve this issue? They make the following assumptions:

- n individuals living for T periods with income $y > 0$; no discounting
- utility $v(0, c)$ without durable, $v(1, c)$ with durable; durable costs B;
- gain from durable $\Delta v(c) \equiv v(1, c) - v(0, c)$
- expected utility from probability α of durable $v(\alpha, c) \equiv \alpha v(1, c) + (1 - \alpha)v(0, c)$

The main logic is similar to our simple model: all n individuals must save separately for t_a periods to purchase the durable.[24] The authors then imagine these individuals forming a

[23] But if participants readjust their savings, then roscas are still welfare improving because with the rosca they save less, and thus enjoy higher consumption – see BCL for details.

[24] Note that this assumes that they find it optimal to save. With non-divisible purchases, it is possible for investors to be discouraged if the the time required to purchase the durable is too long.

rosca with periodicity t_a. Keeping savings and consumption c constant, a random rosca raises utility by the same logic as above. The difference with my simple model is that BCL then let participants choose their optimal level of savings with the rosca. They do so by imagining that rosca participants maximize their ex ante welfare. The authors show that, with the rosca, the optimal level of ex ante saving is *less* than in autarky. This raises welfare since participants consume more but they are still ex ante better off and on average receive the durable faster than by saving on their own. Since this effect operates in all periods, not just for the first rotation, it provides an explanation for the existence of roscas.

The authors discuss bidding roscas in detail and compare them to random roscas. They also discuss roscas between heterogeneous individuals. The last section of the paper is devoted to enforcement issues. The authors note that the voluntary participation constraint is hardest to satisfy for the first recipient of the pot, and argue that social sanctions outside the rosca are necessary to ensure participants continue to pay after receiving the pot.

While this is a nice paper that covers a lot of ground, the authors fail to emphasize that their ex ante welfare optimizing savings rate with a rosca is not time consistent: once a person has received the pot, the optimization problem changes. For instance, if the random order of the first rotation is kept in subsequent rotations, then the time required to wait until the next pot is 10 periods (in my example) and the optimal level of savings should thus be the same as under autarky. Hence the benefit from the rosca disappears.

What this suggests is that a pure economic model is not the right one. People must derive some benefit from participating to a rosca even after the first rotation. One plausible explanation is that roscas help people save. It is like a self-commitment mechanism for people who have difficulties setting money aside, e.g., people who are subject to many pressures to spend whatever money they have – pressures from spouse, children, relatives, but also themselves (impulse purchases, addiction, etc). The beauty of the rosca is that the pot is given away immediately, so that no one ever holds money – and no one can be tempted to abscond with the pot. This feature is particularly attractive if participants are people who find it difficult to hold onto money – but still want to save for a specific target. In my opinion, roscas are best seen as devices designed to help participants solve a self-commitment problem or escaping the pressure of others on their savings. This may explain why roscas are particularly popular among women.

Ambec and Treich (2007) offers a formal model of Roscas as financial agreements to deal with self-control problems. The key ingredient of their model is the utility function:

$$u(y - I_1\ m) + I_1\ S + \beta \sum_{t=2} \delta^{t-1} u(y - I_t m)$$

where:

- y is income – identical across periods
- δ is the discount factor
- m is the cost of a superfluous good yielding an instantaneous utility of $S > 0$
- $I_t = \{0, 1\}$ is the decision whether to purchase the superfluous good at time t

The key is that the superfluous good yields satisfaction S now but the future consumption of m generates no anticipated satisfaction. In each period it is optimal to purchase the superfluous good iff

$$u(y - m) + S > u(y)$$

which we assume. Note that at time t the individual never wishes to purchase the superfluous good in the future – and would like to tie his/her hands to achieve this. A Rosca will make it possible for this individual to enter in a contract that sets money aside to yield a disposable income $\underline{y}$ just low enough so that

$$u(\underline{y} - m) + S \leq u(\underline{y})$$

With this lower income, the individual does not find it optimal to indulge in the superfluous good.

The authors then show that this individual would ex ante prefer a lotery that gives a large amount T with some probability and $\underline{y}$ the rest of the time, to getting y for sure. When they get the lottery, people do purchase the superfluous good m buth they have money left $T - m$ to finance other utility enhancing expenditures. Roscas are a way to implement such a lottery. In this model, it is not even necessary to assume that individuals aim to save for a target durable/non-divisible expenditure.

3.4. Joint liability loans

With the hype surrounding Grameen Bank and other micro-finance NGOs, much hope has been put in micro-credit which, at least for a while, was taken as synonymous with joint liability loans. Such loans have been used by development agencies for a long time. In much of West Africa, for instance, agricultural inputs for cotton are customarily granted to producer groups called *groupements* who are collectively responsible to paying fertilizer and other cotton inputs out of their joint sales. This system has been in use for a long time, perhaps even going back to the colonial period. But the media visibility of joint liability loans begun when Grameen – and many others – began making joint liability loans to Bangladeshi women. Micro-lending to joint liability groups is typically of short duration (e.g., 6 months) and carries high interest rates (e.g., 30-40% a year). Loans are paid in regular installments at short time intervals, e.g., weekly. These installments begin immediately after the loan has been granted, with little or no grace period. A person cannot roll over a loan, i.e., there must be a time interval between two successive loans. Published repayment rates have been high, but many schemes have been heavily subsidized by donors.

It was initially thought that the attractiveness of joint liability loans resides in the fact that repayment rates are higher because, if one borrower cannot pay, the others are asked to pay. As a result, people thought, the risk for the lender is less and this makes lending to the poor if not profitable at least sustainable financially. I do not find this explanation very appealing. Let me illustrate the problem with a simple model.

The joint repayment idea was first formalized by Besley and Coate (1995) – see below for a presentation of that model. To motivate the issues, let me consider a simple setup. Consider a group of size N. Each individual receives a loan L at rate r. Each borrower invests L in an individual-specific project with a probability of failure π. For simplicity, we assume that failure is distributed independently across borrowers. If the project is successful, the borrower can repay; otherwise, repayment is 0. With individual liability, the lender recoups on average $N(1 - \pi)(1 + r)L + N\pi 0$. With joint liability, successful borrowers repay for unsuccessful borrowers. Let us begin by assuming that successful borrowers have deep enough pockets to repay when others cannot. With this assumption, default only occurs if all N borrowers are unsuccessful at the same time. Since success is independently distributed, this arises with probability π^N. The lender can therefore expect to recoup $N(1 - \pi^N)(1 + r)L$, which is clearly greater than $N(1 - \pi)(1 + r)L$ as long as $N > 1$.

This reasoning is predicated on the idea that successful borrowers can pay for unsuccessful ones. This, however, contradicts the idea that borrowers are poor – unless the return on a successful project is extremely large. To illustrate this, imagine that the return to the project is equal to the interest rate r. Further assume that borrowers have no other income or asset than the return from the project – e.g., they only have their labor force and when they work on the project they do not earn a wage w. This means that their income if the project is successful is exactly $w + (1 + r)L$. Finally assume that w is the absolute minimum they need to consume in order to live another day. This means that a successful borrower has just enough to repay ONE loan, not to repay that of others. This implies that joint liability provides no advantage to the lender. If fact it may even be detrimental if all borrowers are punished if one default – in this case, they will all default if one of them cannot pay.

Things would be different if the return on the project is extremely large $>> (1 + r)L$. For instance, if the return is $2(1 + r)L$ then as successful borrower can repay for himself and one other borrower. Note that for this to be true the return on the investment must be at least 100% – possibly much higher if r itself is high. Unless there are reasons to believe that the poor have access to such massively profitable investments, joint liability is unlikely to help raise repayment rates in this fashion. There is in fact little direct evidence that borrowers do in fact pay for each other.

If joint liability is not enforceable because borrowers do not have the money to pay other people's loans, why is joint liability useful? Or is it useful at all?

Two categories of explanations have been proposed by economists to explain why joint liability group loans have been favored by micro-finance organizations: peer monitoring, and self-selection. In case of non-payment by on borrower, it is assumed that other borrowers incur a cost c (possibly smaller than the unpaid loan). This is sometimes presented as $\frac{(1+r)L}{N-1}$, the equal sharing of the unpaid loan. The presence of this cost c may induce co-borrowers to monitor each other to ensure that other bollowers do not take actions that would lead to default. It is not entirely clear what kind of threat they could bring to bear on co-borrowers who misbehave, but it is usually assumed that social sanctions are possible outside the joint liability loan itself. If this is the main benefit from joint liability loans for lenders, then we would expect lenders to encourage people to select co-borrowers among their friends and relatives. In practice, this is seldom if ever the case.

This brings us to the second proposed explanation, best illustrated in the work of Ghatak. Here the focus is on self-selection. So far we have assumed that all borrowers have the same probability of success π. It is however conceivable that they differ. Further assume some homophyly: birds of a feather flock together. Then borrowers with a good project (high π_i) are likely to know other borrowers with a high π_i. If we let borrowers form their own joint liability groups, those with a good project will want to group with other individuals with good projects so as to minimize the chance they have to incur cost c. We should therefore observe assortative matching on project success probability.

This does not solve the lender's problem if the lender cannot distinguish between high π_i and low π_i groups. One could imagine the lender trying to discourage high risk groups from applying for loans. For instance, if borrowers have to share the burden of unpaid loans, c increases with the number of non-performing borrowers, and groups with more high risk borrowers have a higher c. Bad borrowers – those with a bad project – may therefore not form a joint liability group because they expect having to pay for others quite often. And they cannot join a low risk borrower group because those borrowers do not want them. Ghatak expands on this intuition and proposes a separating equilibrium in which incentive-compatible screening contracts are offered that induce high risk and low risk groups to self-select. These contracts require that risky borrowers get charged a higher interest rate but bear less joint responsibility. See below for a short presentation of Ghatak's model.

While this is of course OK from a theoretical point of view, it implies that c can be raised. For instance, if $N/2$ fail to pay, then a performing borrower has to pay back $2(1+r)L$. Unless the return to the project is extremely large – well in excess of 100% – this implicitly assumes that the borrower already had liquid assets or entitlements of the same order of magnitude as L. But then if this is the case, why not do individual loans in the first place, since the borrower could pay in case the project fails. In practice, very poor borrowers typically have hardly any liquid assets, so the upper limit on c is probably reached fairly quickly. If this upper bound is binding, the lender cannot raise c and the self-selection argument loses its value: there will be good groups and bad groups but they will all form and apply for the same kind of loans, and the lender will find it impossible to distinguish among them.

Then there is the issue of joint opportunistic default: co-borrowers may decide to collectively default even when their project is successful. This will be easier if borrowers know each other well and can coordinate their actions. An argument often heard is that this does not happen because if borrowers do not pay, they will not be eligible for another loan. This what disciplines them, it is argued.[25]

There are several 'versions' of this argument. Perhaps the stupidest version is to argue that the size of the loan will increase over time, and it is the hope of obtaining large loans that disciplines people. We have discussed this case in earlier lectures: 'starting small' as a way of weeding out bad risks. But there is a limit to how large the loan can get. Once that size is reached, the disciplining role of the promise of large loans disappears and so does the incentive to repay. This is basically an Ponzi or pyramid scheme. To conclude, a lender who disciplines his borrowers in this manner lives in a fool's paradise: all borrowers will eventually default when they reach their credit ceiling. Something else must be disciplining them.

Another version states that if people fail to pay $(1+r)L$ today, they will not be able to borrow L at the next lending round. But they already have (more than) L today since they are asked to repay $(1+r)L$. Formally the voluntary repayment constraint is:

$$(1+r)L \leq L$$

which cannot be satisfied for any $r > 0$. Hence the promise of a future loan of a similar size cannot induce repayment – and protect against opportunistic default, whether collective or individual.

In a more elaborate version of this argument, Bulow and Rogoff (1989) consider the case in which credit is implicitly bundled with insurance in the sense that no payment is required when the project fails. In this case, they show that IF the borrower can purchase (actuarialy fair) insurance against project failure, then it is still better to default, keep L for the next investment project, and pay the insurance premium with rL. This line of argument would explain the success of joint liability lending (which protects against the risk of unvoluntary defaults and hence serves as insurance) as a substitute for a missing insurance market. Of course this begs the question of (1) why is the insurance market missing and could this be remedied more directly; and (2) why not offer flexible individual loan contracts that incorporate an element of insurance. But as we discussed earlier, all 'real life' loans have an implicit bankruptcy contingency. So what is joint liability adding?

There is another, completely different take on micro-finance that emphasizes NOT standard economic arguments based on imperfect information, but rather behavioral explanations organized around self-commitment. If people have quasi-hyperbolic preferences, they find it difficult

[25] Note that this argument would apply just as well to simple loans.

to save. Peer pressure from the group may serve as an antidote against impulse purchase, or the temptation not to repay. This is a bit like 'alcoholic anynymous' serves as a peer pressure group to surmount the temptation to drink, except that here the group protects against the temptation to spend the money instead of repaying the loan.

Joint liability lending requires the formation of groups, that meet on a regular basis to pay loan installments and attend various 'awareness' meetings. Perhaps a similar outcome could be obtained by keeping the groups and the meetings but omitting joint liability – see Ashaf, Karlan and Yin's (2006) experiment in the Philippines. But then if micro-lending serves as a commitment device, are there alternative commitment devices that have a lower cost? Micro-lending is costly for lenders because of the complication of organizing meetings and of keeping the groups together. Could we do away with the meetings? This is open for research.

This discussion also brings to light the idea that micro-credit is not successful so much because it is CREDIT but rather because it enables people to SAVE. This is not a ridiculous idea. For instance, many people buy a house precisely with this objective in mind: to create a self-commitment device to save for their old days. A home is a good basis for a self-commitment device because people get attached to their home and if they don't pay they get kicked out of the place where they live.

In the case of micro-credit, the horizon of the loan is very short – a few months – the loan amount is very small, and the interest rate is very high. It would be very difficult to imagine a case of such drastic poverty trap that a small loan over six months would take someone out of a poverty trap. Presumably if people can repay their loans in fixed installments – often well before the investment pays off – then they could have saved on their own. The fact that they prefer to borrow at a high interest rate suggests that they cannot save on their own, that they need the discipline imposed by the loan. According to this theory, people enter into debt because they want to save.

The jury is still out as to which model best explains why so many poor people in developing countries have been seduced by joint-liability loans. A number of NGOs now have dropped joint liability from their design and now turn either to group loans without joint liability, to individual loans, or to saving plans (or convex combinations of the above). This, together with the fact that they all insist on repayment starting immediately and following a rigorous schedule at short time intervals, is most consistent with behavioral models. The poverty trap is one that comes from human nature.

3.4.1. Besley-Coate (1995)

Assumptions:

- each borrower has investment project of size 1 lasting one period with uncertain return $\theta \in [\underline{\theta}, \overline{\theta}]$ with distribution $F(\theta)$
- borrower is risk neutral

Individual lending:

- amount to repay is $r > 1$ (note: this is NOT the interest rate)
- bank can impose penalty for default $p(\theta)$ increasing in θ

Repayment constraint:

$$r \leq p(\theta)$$

Let $\phi(r)$ be the threshold return θ at which the repayment constraint is satisfied. The repayment probability (or rate) is:

$$\Pi_I(r) = 1 - F(\phi(r))$$

Further assume that $\phi(1) > \underline{\theta}$. In this case, even if the interest rate is 0, the borrower would not repay if the realized return is low. Hence there is a positive probability of non-repayment for all $r \geq 1$.

Group lending:

- 2 identical borrowers; their realized project returns are θ_1 and θ_2
- they are jointly responsible for the loan: repayment due is $2r$
- bank can impose penalties $p(\theta_1)$ and $p(\theta_2)$ if $2r$ is not repaid

The author examine in detail the repayment game between the two borrowers.[26] They show that under joint liability, the loan is repaid if at least one borrower receives $\theta_i > \phi(2r)$. This is because, by the notation, this borrower prefers repaying the two loans rather than incurring the penalty. If $\phi(r) < \theta_i < \phi(2r)$ for $i = \{1, 2\}$, the loan may be repaid. This means that there exist a Nash equilibrium in the ex post game among borrowers such that their best response is to contribute to the repayment if the other contributes as well. There is also a Nash equilibrium in which they defect. Otherwise the loan is not repaid.

The authors then calculate the probability of repayment $\Pi_G(r)$, assuming that if $\phi(r) < \theta_i < \phi(2r)$ repayment takes place. They obtain an expression of the form:

$$\begin{aligned} \Pi_G(r) - \Pi_I(r) &= F(\phi(r))[1 - F(\phi(2r)] - \\ &\quad [F(\phi(2r) - F(\phi(r))]F(\phi(r)) \end{aligned}$$

The first term favors group lending – this is the additional repayment when one borrower has a return of $2r$ and the other does not pay. The second term goes against it; it corresponds to the case where one borrower has a return above r – and would pay under individual borrowing – but the other has a lower return and they both defect. This means that the above expression is ambiguous – it depends on the distribution of θ. For group lending to dominate it is necessary that r be low enough.

The authors then introduce social sanctions that borrowers can impose on each other. They show that this increases the probability of repayment to situations in which one borrower has a return $\phi(r) < \theta_i < \phi(2r)$ and the other has a return $\theta_j < \phi(r)$ but $p(\theta_j) + s(p(\theta_i) - r, \theta_j) > r$ where $s()$ is the social sanction function. The authors then show that, in this case, repayment is larger under group lending for a large set of parameter values as long as $\phi(2r) < \overline{\theta}$.

Discussion In my simple example, the return on the investment only took two values – success or failure. Here the authors consider returns distributed continuously over an interval. This is of course more general. Note, however, the important role played by $\phi(2r)$ – this is the case when one borrower has enough income to pay both loans. I have argued that this is an unlikely outcome when we talk of very poor borrowers investing in simple things (e.g., chicken). If $F(\phi(2r)$ is very small ($2r$ return very unlikely) then group lending provides LESS repayment than individual liability because borrowers default even when they can pay but the other cannot.

[26] As far as I can tell, they only consider Nash equilibria in pure strategies.

In order to make group lending attractive for the lender, Besley and Coate have to introduce social sanctions. There are two ways to think about this. One is that these are actual sanctions, which is what the authors do. I am not entirely convinced that social sanctions would be found ethical if they are used to punish someone who refuses to pay a bank because he is short of money. Hence I am not entirely convinced that social sanctions of this kind would be very effective. Another is to think of $s(.)$ more generally as an additional reason to repay when in a group, e.g., that being in a group resolves self-commitment problems and reduces pressure to divert resources to friends and family members.

3.4.2. Ghatak (2000)

Assumptions:

- each borrower with a project of size 1; project can fail or succeed: $x \in \{S, F\}$
- two types of borrowers, 'risky' and 'safe' with probability of success $0 < p_r < p_s < 1$ in proportion θ and $1 - \theta$
- project outcomes are independently distributed
- return to project R_i if successful and 0 if fails
- borrowers are risk neutral with reservation payoff $\overline{u}$
- opportunity cost of capital is $\rho \geq 1$

Individual lending: Assumptions:

- borrower does not pay if project fails
- borrower pays in full if project succeed – this is an assumption; not voluntary repayment

With full information, the bank charges interest rate:

$$r_i^* = \frac{\rho}{p_i} \text{ for } i = r, s$$

This shows that the interest charged is higher for risky borrowers, those with a lower probability of success.

If the bank cannot tell the type of the borrower, things get more complicated. The author discusses various configurations and shows that underinvestment is possible in a pooling equilibrium where the lender cannot distinguish between borrowers.

Group lending Assumptions:

- borrowers must pay c on top of r per member of the group who defaults; this is assumed possible

With this assumption, Ghatak shows that individuals sort into high risk and low risk groups. He then shows that this generates groups that satisfy the single-crossing property. This is the clever part because this property is needed in adverse selection models to ensure incentive-compatibility in screening different types of agents. With this property satisfied, lenders can devise loan contracts that induce groups of risky borrowers to separate themselves from safe borrowers.

The optimal separating joint liability contracts (r_s, c_s) and (r_r, c_r) have the property that $r_s < r_r$ and $c_s > c_r$: risky groups are charged a higher interest rate but a lower contribution in case of default. Ghatak does not discuss enforcement constraints, so it is unclear what conditions are required for c_s and c_r to be paid. See Besley and Coate for a discussion of strategic repayment issues in joint liability loans.

It is also unclear how much repayment lenders can achieve with group lending: in the paper, contributions c are set in a way that is largely if not entirely disconnected from the actual amount due r and the size of the group. Joint liability normally means that c should be $\frac{r}{n-1}$ if one borrower defaults, where n is the size of the group. This would be the contribution that is required to ensure that all the contractual liabilities of the group are met. If m borrowers default, then we should have $c = \frac{r}{n-m}$, which implies that under joint liability c varies with the number of people defaulting, and hence it is not a contractual constant, as assumed by Ghatak. But this assumption is a reasonable approximation if the group is large and default does not occur too often. If we only have to worry about a single default, a larger c for safe borrowers means a smaller group size for them. This is an easily implementable condition.