Access to Finance
Chapter 2, Handbook of Development Economics, Volume 5
Dani Rodrik and Mark Rosenzweig, eds.

Dean Karlan & Jonathan Morduch

June 2009

Contributions to this research made by a member of The Financial Access Initiative and Innovations for Poverty Action.

The Financial Access Initiative is a consortium of researchers at New York University, Harvard, Yale and Innovations for Poverty Action.

NYU Wagner Graduate School
295 Lafayette Street, 2nd Floor
New York, NY 10012-9604

T: 212.998.7523
F: 212.995.4162
E: contact@financialaccess.org

www.financialaccess.org

Innovations for Poverty Action applies rigorous research techniques to develop and test solutions to real-world problems faced by the poor in developing countries.

Innovations for Poverty Action
85 Willow St, Building B, 2nd Floor
New Haven, CT 06511

T: 203.772.2216
F: 203.772.2428
E: contact@poverty-action.org

www.poverty-action.org
Access to Finance

Dean Karlan and Jonathan Morduch

Chapter 2, Handbook of Development Economics, Volume 5
Dani Rodrik and Mark Rosenzweig, eds.

Abstract

Expanding access to financial services holds the promise to help reduce poverty and spur economic development. But, as a practical matter, commercial banks have faced challenges expanding access to poor and low-income households in developing economies, and nonprofits have had limited reach. We review recent innovations that are improving the quantity and quality of financial access. They are taking possibilities well beyond early models centered on providing “microcredit” for small business investment. We focus on new credit mechanisms and devices that help households manage cash flows, save, and cope with risk. Our eye is on contract designs, product innovations, regulatory policy, and ultimately economic and social impacts. We relate the innovations and empirical evidence to theoretical ideas, drawing links in particular to new work in behavioral economics and to randomized evaluation methods.

Keywords: Microfinance, Microcredit, Microinsurance, Credit, Savings, Insurance, Behavioral economics, financial intermediation, economic growth, randomized controlled trials.

JEL Codes: O16, O17, G21, G22, D03, I32

1. Introduction
2. Global financial access
3. Financial intermediation and economic growth
4. Returns to capital
5. Credit market innovations
6. Understanding saving
7. Insurance and risk management
8. Governments, businesses, and non-profit institutions
9. Conclusions

1 Dean Karlan: Yale University, dean.karlan@yale.edu. Jonathan Morduch: New York University, jonathan.morduch@nyu.edu
1. Introduction
Many interventions have been proposed to solve entrenched development problems, or at least to make noticeable dents in poverty levels. The list of accumulated hopes is long, including better nutrition to catapult levels of productivity and wages; control of population growth to free resources for human capital investment; education for girls to fight inequalities and bring empowerment; and stronger property rights to unleash markets. Each hope is grounded in good reason, and each intervention holds a place in the larger scheme of development strategies. But none on its own has proved to be a catalyst on the scale imagined by its chief proponents.

In recent years, much hope has been placed on the transformative power of financial access. It is, in many ways, the boldest claim so far, and the most unlikely. The best-known advocate has been Muhammad Yunus, the co-winner of the 2006 Nobel Peace Prize alongside Grameen Bank, the bank Yunus founded to serve the poor of Bangladesh. Yunus speaks eloquently and forcefully about the power of access to small loans—dubbed microcredit—to transform the businesses of poor households.2 With those loans, Yunus argues, incomes will grow and, with rising incomes, children will be given long-denied opportunities. As Yunus (2006) declared in his Nobel lecture in Oslo: “we are creating a completely new generation that will be well equipped to take their children out of the reach of poverty.”

Yunus’s argument has grounding in economic theory. The argument aligns with explorations of credit rationing that show that when lenders lack good information on customers and contracts are costly to enforce, outcomes are not necessarily Pareto efficient (Besley, 1994; Stiglitz and Weiss, 1981). Innovations in credit markets can thus, in principle, bring gains in both efficiency and equity. The common assumption that the marginal return to capital is large when capital is scarce reinforces the claim that the “unbanked” poor have sizeable returns to reap from financial access.

Yet, as a practical matter, commercial banks have had difficulty providing such access profitably. The unbanked (and under-banked) tend to be poor and often lack titled assets to offer as security for loans. Moreover, many of the unbanked want to make transactions at too small a scale to attract much interest from profit-seeking institutions (Cull, et al., 2009b; Johnston and Morduch, 2008). The “microfinance revolution” has thus had to contend with incentive problems alongside more prosaic challenges imposed by transactions costs. More fundamentally, the list of other factors correlated with poverty is long (including low education levels, poor health, discrimination, and weak labor markets), and these challenges risk undermining the effectiveness of financial access in raising incomes. The evidence to date shows that access to capital may be powerful for some, but it does not yield high returns for all.

2 The term microcredit was later expanded to “microfinance” by Yunus’ followers to be inclusive of other financial services for the poor, including savings and insurance.
This way of thinking, centered on the productive potential of capital, requires scrutiny in part because the evidence so far is mixed. But, more importantly, the conceptual frame centered on providing small loans for small businesses is too limiting. While Yunus’s vision of unleashing the productive potential of millions of small-scale entrepreneurs yields a powerful narrative, it risks blinding policymakers and practitioners (and researchers) to the broader financial needs of poor and low-income households. Those broader financial needs are, in many ways, similar to the needs of richer households: mechanisms to manage cash flows, devices for accumulating assets in both the short and long-term, and tools for coping with risk. Access to capital to expand businesses can generate income that facilitates these tasks. But as Collins et al (2009) show through year-long “financial diaries” that track the financial lives of poor and “near poor” households in Bangladesh, India, and South Africa, financial activities are most often driven by a basic set of needs—e.g., to get food on the table every day, deal with illness, pay school fees and other sizeable expenses, and seize investment opportunities as they arise. None of these needs is necessarily tied to running small businesses, and all are as important for employed people in cities as they are for village women running micro-enterprises.

Work on “access to finance” is shifting to embrace the idea of providing banking services (credit, savings, and insurance) rather than primarily delivering microcredit for small-scale business. This chapter describes ways of thinking about this transformation, with an eye on innovations that help expand and improve financial access in poor communities. We focus sharply on contract designs, product innovations, and regulatory policy, complementing earlier surveys on access to finance, including *Handbook of Development Economics* chapters written by Gersovitz (1988) and Besley (1995), which, to a far greater degree, focus on theories of banking, macroeconomic frameworks, and informal finance. While most poor and low-income households continue to conduct most financial transactions through informal mechanisms, our focus on informal finance is largely instrumental. This is not to diminish the importance of the informal sector. But given the parameters of the chapter, we look to informal mechanisms mainly as guides for product design, and contract possibilities, and for context in understanding the measured impact of specific innovations.

*Mechanisms Matter*

---

3 We do not focus on remittances in this chapter, except briefly in the technology section as it relates to interventions such as electronic banking. However, we note the growing role of remittances in international capital flows (and in poor households’ financial lives). The World Bank reports that remittance flows to developing countries totaled $251 billion in 2007, more than doubling since 2002 (Ratha, et al., 2008). Part of the increase in remittances may be due to improved data recording and the depreciation of the US dollar (Savage and Harvey, 2007), but, even so, the figures likely underestimate total flows (Orozco, 2007). In particular, these data only capture officially recorded transfers, and a large portion of remittances appear to flow through informal channels and thus remain unrecorded. Estimates that impute unrecorded flows yield levels of remittances that are greater than foreign direct investment flows and more than twice the level of official aid received by developing countries (Mohapatra, et al., 2006).
For all of the unknowns, we have acquired one central understanding about financial behavior: that mechanisms matter. Old debates—about whether the poor can repay loans reliably, or whether they can pay high interest rates, or save, or insure—need to be re-cast. Yunus’s fundamental insight was to show that the poor are bankable if the right lending mechanism is used. The earliest mechanism to gain attention is the “group lending” contract, in which neighbors meet together to take loans and collectively assume responsibility for their repayment, mitigating problems imposed by information asymmetries and costly external contract enforcement. New evidence shows that this is only one of several key mechanisms, and probably not the most important (Armendariz de Aghion and Morduch, 2000 and; 2010; Giné and Karlan, 2009). But no matter their individual roles, when taken together the successes of the financial mechanisms have changed the terms of debates. Grameen Bank (like other microfinance institutions) reports loan repayment rates above 98 percent despite lending to poor households, most of whom lack collateral and experience with banks. Policymakers, “social” investors, and academic researchers have taken note.

Similar understandings of the importance of mechanisms have emerged with regard to saving, and, to an extent, insurance. For example, studies show that poor households often seek specific, structured financial tools to achieve their savings goals—not just generic saving accounts (Ashraf, et al., 2006b; Collins, et al., 2009; Duflo, et al., 2006). Old prejudices held that poor households lack the surpluses to save much (Bhaduri, 1973). The idea conformed to a notion of poverty defined as having income that falls below a minimal threshold necessary for basic subsistence. The logic holds that if you are struggling to meet your needs today, saving up will inevitably be a slow process. Yunus’s initial push for microcredit (rather than “micro-saving” or “microfinance”) thus made sense as a way to speed the process of transformation. But by the 1990’s, the view that the very poor are unable to save has been turned back, prompted in particular by lessons on wide-scale saving in Indonesia (Patten and Rosengard, 1991). As Banerjee and Duflo (2007) show in their analysis of tens of thousands of households in thirteen developing countries, even the very poor, living on under $1 a day per person, spend relatively heavily on what appear to be non-pressing expenditures (like social and religious expenditures). Against that background, it is easier to see room for saving within the budgets of the poor, and Grameen Bank itself has now introduced an array of flexible and structured saving products.

Insights follow as well from paying closer attention to the psychology of financial decision-making (e.g., Thaler, 1990). Lessons from behavioral economics are naturally relevant for choices by poor and low-income households, and we see applications in research on pricing, saving, insurance, and debt traps. The new work shows the roles of limited self-control, loss aversion, and mental accounting and their implications for product design and marketing (Bertrand, et al., 2006). While we are far from having a grand, unified theory based on the psychology of financial choices, the evidence so far demonstrates gains from expanding beyond (but not away from) traditional economic intuition—in which the way that products and choices are presented to consumers is essentially immaterial. The empirical evidence is mounting that both product design features and presentation can matter greatly.
Testing What Works

From a macro perspective, expanding financial access holds the promise of increasing economic growth by spurring investment in under-funded enterprise, following the logic of Gurley and Shaw (1955) and McKinnon (1973). On the savings side, expanding access to reliable, low-cost deposit accounts promises to increase the capital stock. Given that the expansion of access favors lower-income populations, these steps also promise to reduce poverty and inequality. All of this is true in principle, but there is little evidence so far that expanding financial access through microfinance has had notable macro impacts anywhere. Only in a few countries—Bangladesh and perhaps Indonesia and Bolivia—is the scale of microfinance large enough to even imagine the possibility. We do know from cross-country evidence that financial deepening correlates with inequality reduction (Demirgüç-Kunt and Levine, 2008), but the lack of scale means we have no firm results with regard to microfinance specifically, and endogeneity and sample size issues hamper causal inference in cross-country regressions.

Much of the action has thus been at the micro level, and turning there requires a different set of lenses. The macro perspective puts a natural focus on savings primarily as a way to increase wealth and borrowing as a way to fuel investment. But for households, borrowing is also an important way to cope with emergencies and to pay for household and social expenses. To this extent, borrowing is welfare enhancing if not always output-increasing. Saving too can be an important way to smooth consumption from month to month and to cope with within-year expenses, and not chiefly a means to build up long-term balances (Rosenzweig, 2001). Again, saving may be welfare enhancing even if not particularly output-increasing. In their close look at the financial lives of poor households, for example, Collins et al (2009) find a common pattern of intensive use of saving instruments but relatively small average balances. Turning to the micro level also gives risk mitigation a prominent place in expanding financial access, and we review the growing movement to provide micro-insurance. Without much formal or informal insurance, borrowing (whether at zero interest from neighbors and relatives or at high prices from moneylenders) becomes by default a primary way to cope with emergencies.

The evidence so far suggests that financial access will not, on its own, be enough to take children out of the reach of poverty on a massive scale. Nor does the evidence suggest that finance alone is necessarily as powerful as finance coupled with other interventions – like training and healthcare. But the most striking conclusion from the available evidence is in fact that many of the big questions are left unanswered. There have been few fully convincing studies of impacts, and little rigorous investigation of whether the very poor can benefit from financial access to the same degree as the less poor—or perhaps whether the very poor will benefit more than others. Either possibility is consistent with economic theory and is at root an empirical issue. Similarly, the knowledge of saving behaviors and risk management strategies of the poor is only now accumulating, as is our understanding of price sensitivity and the demand for particular qualities of service.
Establishing appropriate counter-factuals is a critical challenge for researchers. Convincingly teasing apart the roles of mechanisms and their impacts on customers has been slow-going, though progress is now being made through the adoption of approaches to evaluation that incorporate experimental elements, most importantly randomized controlled trials of various kinds. The new approaches draw on decades of experience with evaluations of medical treatments, and represent the best ways developed yet to address the selection biases and omitted variable biases that undermine the credibility of evaluations.\(^4\)

The potential biases are particularly acute when assessing financial interventions. Microfinance customers tend to be especially entrepreneurial and energetic relative to their non-participating neighbors. This causes self-selection issues, which make non-experimental evaluations challenging—and which tend to bias estimates of impact to over-state actual benefits. Even if self-selection is not an issue, financial institutions are apt to carefully screen potential customers, filtering the pool to find the most promising customers, and seeking the most promising locations in which to operate. Again the biases tend to lead to over-stating actual benefits if banks target as described above. Alternative selection processes can, correspondingly, lead to underestimates of impact.

Randomized controlled trials can eliminate the resulting selection biases by building evaluation methods into program design. Recent methodological innovations in experimental design aim to ensure that evaluations are cost-effective, ethically appealing, and useful for the programs and customers. But experimental approaches have limits (many of which are shared with non-experimental modes of evaluation), and have only picked up steam in the past five years in their application to issues around financial access.

To form future policies wisely, randomized controlled trials should be pushed in two directions: first, researchers need to replicate studies in different settings. Learning that a given approach to microfinance worked in one place, with one institution, at one point in time is not sufficient to know what to do in the future. What works in Bangladesh may not work in Argentina. What works in the city may not work in a small town or a village. The problem of “external validity” (i.e., uncertainty around how far it is appropriate to generalize a particular study to other contexts) is an old but often ignored problem in applying empirical research to policy decisions. Replicating studies allows analysts to begin to address the problem of external validity by building a clear understanding of the necessary context for an intervention to work (i.e., robust tests of theories that account for how the context will influence the outcome). It may be that, on a macro level, certain interventions work best in boom economies, but not in low-growth scenarios. Or, on an individual level, certain interventions may, say, work best with women who have little to no preexisting power in the household, while having little impact on women who already exhibit considerable control within their families. Replication, combined with attention

\(^4\) Randomized trials have proved especially valuable when evaluating the working of new mechanisms, a focus of the present chapter. Other econometric approaches have been (and will continue to be) valuable in assessing directions for financial access and relationships between economic and financial variables.
to theoretical relationships, can help us understand the underlying failures and the contexts in which innovations succeed.

The second direction involves learning why things work. We ask this on two levels. The first concerns market structure. In order to think about whether an idea has promise for solving a market failure (as is the claim for microfinance), it is necessary to understand why the market failed in the first place, and how this intervention was able to solve the specific failure. The second level is micro, regarding individual decision-making. Here, we need to understand better the mechanics of choice, particularly for the choices faced by poor households. The information allows us to predict how the choices made by low-income households (and the outcome that can be achieved) will be affected by changing the available financial options and tools, including pricing. The aim is to learn information that is forward-looking, rather than confining efforts to only looking backward to assess the impacts of existing interventions.

Clearer data on impacts, market structure, and household-level (and often individual-level) decision-making are critical for weighing major public policy issues—and are necessary complements to ethnographic, financial, and administrative data. The most voluble debates concern the appropriate use of subsidies and the setting of price regulations for financial institutions serving poor households. But a broader set of concerns has received less systematic attention: whether investing in the sector as a whole is the most cost-effective way for donors to achieve their missions, relative to alternative interventions that reach poorer households, reach larger businesses, or that focus on interventions like health, education, and infrastructure. The microfinance movement has proved the possibility of creating viable economic institutions on a large scale, and the challenge now is to more carefully assess social and economic impacts.

The next section describes what we know about the gaps and accomplishments in providing financial access globally. Section 3 reviews the links between financial intermediation and economic growth, drawing on both theoretical and empirical work. Much of that hinges on assumptions about the returns on assets in small businesses, the topic of Section 4. Section 5 turns to credit market innovations, and section 6 to savings. Section 7 then describes emerging work on risk management and insurance. Section 8 focuses on the policy landscape and the roles of governments, businesses, and non-profit institutions. The final section draws conclusions.
2. Global Financial Access

Gaps in financial access remain stark. Using survey data combined with aggregate indicators Demirgüç-Kunt et al (2007) report estimates of the share of populations with accounts in formal and semi-formal (e.g., microfinance) financial institutions. More than 80 percent of households in most of Western Europe and North America have an account with a financial institution. In Central Asia and Eastern Europe 60 to 80 percent are estimated to have accounts, with Latin America exhibiting variation ranging from less than 20 percent in Nicaragua to more than 60 percent in Chile. Estimated usage in Asian countries generally ranges from 40 to 60 percent. A World Bank study in rural India, for example, finds that about 40 percent of households have deposit accounts, 20 percent have outstanding loans, and only 15 percent report having any insurance (Basu, 2006). In much of Sub-Saharan Africa, fewer than 20 percent have accounts. Only in Botswana, the Gambia, and South Africa are the estimates above 60 percent (Demirgüç-Kunt, et al., 2008).

Taken together, the results suggest that the number of “unbanked” and “under-banked” adults worldwide could be two to three billion people: precise figures have not been aggregated. Against that backdrop, the rapid expansion of microfinance has been stunning but still leaves substantial gaps. The Microcredit Summit Campaign Report of 2007 reports a growth of 885 percent in the number of clients from 1997 to 2006 – an average annual growth rate of 29 percent per year. In 2006, 3,316 institutions reported to the organization, and those institutions reached 133 million clients; a year later, the number had swelled to 154 million. In 1997, only 618 institutions were found, cumulatively serving 13.5 million clients; remarkable growth considering that 93 million of the 133 million at the end of 2007 are judged to be among the “poorest,” an income segment that traditional banking institutions have long considered unbankable. Of the poorest microfinance counted in the survey, 90 percent are in Asia, mostly in Bangladesh and India (Daley-Harris, 2009). Overall, most microfinance customers are found in Bangladesh and India, with the next largest group in East Asia and the Pacific.

Still, even in Bangladesh, there are substantial gaps in financial access. The number of loans per 100,000 people in Bangladesh, for example, is 54.73 and the number of deposits for 1,000 people is 228.75 placing it at the 31st and 43rd spot respectively in a World Bank survey of 53 developed and developing countries (Beck, et al., 2006). Without the spread of microfinance institutions, Bangladesh would have ranked considerably worse, but the numbers show that there is further to go in spreading access.

It is not just availability that matters. Fees, costs, and documentation requirements also serve to limit financial access. Beck, Demirguc-Kunt and Martinez Peria (2006) report on an important survey of the largest commercial banks in a large sample of countries, documenting price and non-price barriers associated with deposit, credit and payment services. The survey shows critical variations across countries in the degree of physical

---

5 The “poorest” is defined as being below the World Bank UN international poverty line of $1 a day in purchasing power adjusted dollars, or being in the bottom half of a country’s poor population (as defined by the local poverty line).
access to formal financial institutions, documents required to maintain accounts, and
costs (e.g., minimum balance requirements and fees). In one dramatic example, they find
that opening a checking account in a commercial bank in Cameroon required a minimum
deposit of over $700 (a figure greater than Cameroon’s GDP per capita). In Sierra
Leone, maintaining a checking account required annual fees exceeding 25 percent of
Sierra Leone’s GDP per capita. Getting a small business loan processed in Bangladesh,
Pakistan, or the Philippines can take over a month. And transferring $250 dollars abroad
cost $50 in the Dominican Republic. These extreme examples are made more striking by
the fact that other banks have managed to drop minimum balance requirements, cut
annual account fees, speed up loan processing, and slash costs for sending remittances.

Microfinance has expanded in part due to the rise in foreign capital investment. Between
2004 and 2006, foreign capital investment in microfinance tripled to $4 billion; by 2007,
investment had reached $5.4 billion. Institutional investors lending to microfinance
institutions reached US$550 million in 2006. The majority of the global capital flows go
to about 30 countries in three regions, though: Latin America, Eastern Europe and
Central Asia. Africa and Asia receive only 6 to 7 percent of foreign investment (Forster
and Reillie, 2008). The mismatch of capital flows and the locations with a greater
prevalence of poverty is startling given the emphasis by microfinance leaders on poverty
reduction, but investors have been wary about the perceived lack of management capacity
and regulation that imposes hurdles.

Despite the capital flows from social and commercial investors, the greatest microfinance
outreach at this juncture is not from commercial institutions but from public sector banks,
non-governmental organizations, and self-help groups (an Indian hybrid based on
partnerships between NGOs and banks; Cull, (2008). One report reviewed 2,600
microfinance institutions to better understand the institutional landscape (Gonzalez and
Rosenberg, 2006). They found that non-governmental organizations served 25 percent of
the 94 million borrowers found in 2004, with self-help groups serving 29 percent.
Commercially based microfinance banks and licensed “non-bank financial institutions”
 served only 17 percent, though the composition is shifting toward commercial players,
pushed by the transformation of non-governmental organizations into non-bank financial
institutions.

These shifts are likely to affect the nature of services delivered to customers. Cull et al.
(2009b) use data from the Microfinance Information Exchange (MIX) to analyze lending
models and outreach of 346 leading microfinance institutions serving 18 million active
clients in 2002-2004. They find that two-thirds of commercially oriented microfinance
banks lent through individual methods (i.e., standard bilateral loan contracts), while
three-quarters of nongovernment organizations used group-lending methods in the
original spirit of the Grameen Bank. The latter tend to target poorer households and often
use the groups for social support, while the individual lenders tend to target “upmarket”
clients looking for larger loans.

This broad picture of financial access is starting to gain detail, but it remains too
imprecise to guide local policy. Details that may seem trivial—how a survey question is
phrased, for instance—turn out to strongly shape responses (Cull and Scott, 2009). More generally, large, one-time surveys tend to miss important information, partly because respondents hesitate to disclose intimate information about their financial lives to outsiders, especially about informal activities. Such discrepancies are revealed by the collection of “financial diaries.” In an intensive data collection effort, Collins et al. (2009, figure A1.1) collected information on all household financial inflows and outflows for small samples in Bangladesh, India, and South Africa, repeating the interviews every two weeks for a year. The initial interviews in South Africa greatly under-counted inflows, a deficit that was narrowed to within 6 percent only after about 6 meetings (i.e., 3 months of repeated interviews). Much of what was under-counted was informal. Savings clubs, reciprocal credit arrangements with friends and family, and other informal financial mechanisms turned out to be abundant, but seldom picked up by large one-time surveys of the sort collected by government agencies and research organizations. The 42 households in the Bangladesh sample reported using 33 different devices, with no household using fewer than 10, while two thirds percent of South African diarists belonged to at least one informal savings club. In all three countries, informal mechanisms were used more frequently by the poor than any other kind to form lump sums of money, even in the South African sample where many respondents held bank accounts.

3. Financial intermediation and economic growth

Economists have long linked the expansion of financial markets to the spread of broader economic activity. By the same token, economists have focused on ways that barriers to financial markets undermine economic efficiency. In the 1970s, economists turned their focus on regulations in many countries that capped interest rates on loans. Interest rates serve many roles, and one is to screen the quality of investments. When interest rates are set artificially low, borrowers are undeterred in investing in businesses that have relatively low returns. Artificially low interest rates also lead to excess demand for credit—and thus, inevitably, to credit rationing. Goldsmith (1969) stitched together these pieces of analysis to argue that interest rate caps undermine the average quality of investment, yielding “financial repression”. The notion of financial repression was extended by McKinnon (1973) and Shaw (1973) who turned to savings, focusing on the ways that interest rate caps ultimately reduced returns on saving as well, ultimately reducing both the quality and quantity of investment. The McKinnon-Shaw treatises drove broad arguments for financial liberalization (a push, notably, to allow interest rates to rise to levels determined by markets), and their ideas fueled a specific assault on rural credit “directed lending” programs, led by researchers associated with the Ohio State University Rural Finance Program (e.g., Adams, et al., 1984).

The association of financial expansion and economic growth is well-established in the empirical literature. The causal link is harder to establish, however, since economic growth spurs financial expansion just as financial expansion can spur growth. Levine (2005) reviews the basic empirical associations, arguing that the link from finance to economic growth cannot be explained merely by reverse causation (drawing on cross-country regression analyses including those by Rajan and Zingales, 1998; and Beck, et
al., 2000). These empirical findings are based on data aggregated at a country level. The empirical linkages cannot be tied to the expansion of financial access by households (as opposed to firms), nor to the spread of microfinance. At this date, the penetration of microfinance is too low in most countries to draw reasonable inferences about broad economic impacts (Honohan, 2008). Indeed, the challenge at this point is to establish basic household-level impacts of microfinance.

A related strand of cross-country literature, though, turns to the distributional impacts of financial expansion. It does not ask about the impact of financial access by the poor on macro indicators, but instead asks about the impacts of financial deepening on poverty and inequality. The impacts are set out in the theoretical model of Loury (1981), for example. The focus of his model is the intergenerational transmission of inequality; parents’ inability to borrow to fund investment in their children’s human capital means that inequality of resources in a given generation translates into inequality in the next generation. In Loury’s model, redistribution can thus improve economic efficiency. While it is not a stress of the paper, relaxing borrowing constraints will also improve efficiency, as well as reducing inequality and its persistence over time. The basic result— that borrowing constraints reduce efficiency and exacerbate inequality by diverting capital from low-income households with high-return investments—emerges in a string of more recent theoretical papers, including Galor and Zeira (1993); Aghion and Bolton (1997); and Banerjee and Newman (1993). Greenwood and Jovanovic (1990) build a model in which financial development can increase inequality as better-off households are, at first, best positioned to take advantage of finance. The logic follows from Townsend (1978; 1983) who builds off the idea that investment in creating financial systems is costly. In Greenwood and Jovanovic (1990), richer segments of the population thus invest in financial infrastructure first; over time a broader swath of the economy benefits, so that inequality widens then narrows with financial development. In Greenwood-Jovanovic (1990), though, financial deepening is poverty-reducing at all points.

As with the empirical literature on economic growth, directions of causality are difficult to establish. Demirgüç-Kunt and Levine (2008, p.1) are left to note that economic researchers have done “an inadequate job of examining how formal financial systems affect the poor. We find this surprising because many of the profession’s most influential theories on intergenerational income dynamics advertize the central role of financial market imperfections in shaping the economic opportunities of the poor.” Clarke et al. (2006) and Beck et al. (2007) are among the few papers to investigate the link across countries. Both focus on the role of private credit on measures of inequality. The private credit variable captures the value of credit offered by financial intermediaries (excluding the central bank and state-owned development banks) to the private sector as a fraction of GDP. Clarke et al find that financial development is associated with inequality reduction in a data set for 83 countries in 1960-95, in line with Galor and Zeira (1993) and Banerjee and Newman (1993). The result is robust to instrumental variables estimation using the origin of the country’s legal system as a determinant of the degree of financial development—under the assumption that historical origins play no current role in explaining outcomes once contemporaneous variables are included in specifications.
Beck et al (2007) provide similar results, extending the analysis by adding countries and years, taking the number of observations from 170 to 245. They rely on the timing of trends to make causal claims, given a dearth of alternative, credible instrumental variables. Their main conclusion is that financial development “disproportionately boosts incomes of the poorest quintile” and thus reduces income inequality. Financial development is associated with a reduction in the population share living on less than $1 a day as well. Most of the long-run gain made by the poorest fifth (60 percent) comes from general growth effects and the balance (40%) results from reductions in income inequality. The broad conclusion is that financial development is good for the poor—though, here, the link occurs mainly through trickle-down effects.

One of the reasons that so little work has been done to tackle these kinds of links in cross-country data rests with the lack so far of breakthrough empirical approaches to solve statistical identification problems. In contrast, there has been a great growth of micro studies that attempt to link financial access to household well-being and decision-making. The micro studies have the advantage of isolating the impacts of particular kinds of financial intermediation (rather than focusing on financial development broadly measured in an economy). The application of general equilibrium models calibrated to specific economies also holds promise as a way to integrate micro and macro analyses (see, for example, Townsend and Ueda, 2006 calibration of a model with fixed financial costs to Thai data, 1976-96). The micro studies, in addition, hold the promise of evaluating specific assumptions underlying theoretical models, such as the non-convex production technologies that undergird models like those of Galor and Zeira (1993) and Banerjee and Newman (1993)—a research program outlined by Banerjee and Duflo (2005).
4. Returns to Capital

If there is one fundamental argument in the global microfinance movement, it boils down to beliefs about patterns of returns to capital. On one hand is the belief that poor households can earn higher returns than richer households. The idea stems from the assumption that poorer households are more likely to face binding financing constraints and thus will get an especially big boost in productivity from access to finance. The other side argues that this logic holds only to a point: the very poorest households likely lack the wherewithal to be reliable bank customers and are better off being served by other economic and social interventions (like education and health services that build human capital). We argue below that the terms of that frame are too stark and that generalizations based on income level alone conceal as much as they reveal. All the poor are not alike. More interesting questions surround (1) How to identify non-income dimensions along which patterns of returns can be differentiated and (2) how to identify other interventions (e.g., financial literacy, skills training, marketing, health) that may raise returns to capital for low-income populations.

Theoretical models that yield credit constraints usually depict efficiency gains from expanded financial access; relaxing constraints means that the productive potential of entrepreneurs is unleashed—farmers who lack the cash to buy enough fertilizer at planting time, weavers who can’t buy sufficient yarn, shopkeepers who can’t adequately build their inventory (e.g., Banerjee and Newman, 1994). Microcredit advocates like Muhammad Yunus similarly focus on the gains from promoting “micro-enterprise”. Yunus argues that the returns to financial access are bound to be large—large enough in some cases to transform livelihoods and permit sustainable exit from poverty (Yunus, 2006). Even a small bit of extra cash, Yunus argues, can transform money-starved, micro-scale businesses.

The idea gives a place to start, though it ignores the observation that much credit is used for non-business purposes. Still, the idea gives the simplest defense of the claim that poor households can afford the high interest rates often charged by microfinance institutions. Rosenberg (2002) has put forward the claim most sharply in a much-cited publication of the Consultative Group to Assist the Poor. The implication is that poor entrepreneurs can afford high-priced credit (perhaps even better than some richer customers), and that poor entrepreneurs can and should pay the fees required to cover costs, be they 20 percent or 40 percent per year or possibly higher.

The case is backed with anecdotes. Take, for example, the story of Vidalia Mamami, a 43 year-old vegetable seller in Tacna, Peru. She sells vegetables from a stand in a local market, and her earnings help support her husband and five children. She had been in business for 21 years, but only recently turned to Pro Mujer, a microfinance NGO:

With my first loan I was able to buy more merchandise for my business and I was able to add vegetables and condiments, which have increased my earnings. Before, I earned 18 to 20 soles per day by selling only fruit,
but now that I have added vegetables and condiments, I earn an average of 30 to 35 soles per day. This money has allowed my family to eat better and allows me to do things for my children that I could never do before. I remember how my older children were not able to go to school because we didn’t have enough money.\(^6\)

The story puts together a dramatic increase in earnings and ties it to broader social impacts. For economists, it resonates in large part because it aligns with the well-understood model of production under imperfect credit markets. Still, the anecdotes tend to reflect the best cases, and the theoretical analyses assume that constraints bind. The theorist’s job is to focus narrowly, abstracting from other variables that are apt to determine profitability in practice.

The belief that that many poor households are in fact relatively weak prospects for loans, and that they can take better advantage of other interventions (schools, health clinics, savings accounts, insurance, and the like), also makes sense. For those who argue from this side, it is unclear how many “unbanked” entrepreneurs have the skills, business connections, political access, and other inputs that can help in running a truly thriving enterprise; thus their returns to capital may remain low. Second, households with more capital may be able to reap returns to scale unavailable to poorer households. Marguerite Robinson (2001) has, for example, drawn this conclusion in her sweeping assessment of the “microfinance revolution”, and it drives Dale Adams’ wariness of microfinance as a poverty-reduction tool (e.g., Adams and Pischke, 1992). So while microfinance advocates like Muhammad Yunus see credit as a human right (Yunus, 2006), others counter that poorer households may have such low returns that expanding credit access to the poorest might only create a heavy debt burden.

Thus, much of where one stands on ongoing microfinance policy controversies -- Should credit be targeted to the poorest? Are their better interventions for donor dollars? Should interest rates be subsidized for the poorest? Is there a trade-off between financial sustainability and depth of outreach? – is bound up with what one believes about patterns of returns to capital.

Framing the question

Though tempting, these are not questions that can be fully answered by simply looking at whether poor households do pay high interest rates. First, this kind of “market test” gives no sense of the level of gain that households experience. To see the point, consider the case in which microcredit is priced so that loans are only just worth taking. The interest rate, for example, might be 40 percent while the expected return to capital is 45 percent. The 5 percentage point gain is an important incremental gain (and will keep customers coming back for loans), but it is not a transformative change—and not the kind of gains

---

asserted by Yunus. Second, households may be caught in debt traps, paying interest but falling deeper into a hole.

More important, the market test tells us whether some people can pay high interest rates, but it does not tell us anything about people who are not borrowing. Are they not borrowing because they cannot afford to? Or because they have no desire to (but could afford to if necessary)? Household surveys that look at a broad population are needed to see the bigger picture.

Johnston and Morduch (2008) show how this matters with evidence from a survey in which loan officers employed by Bank Rakyat Indonesia, a pioneering microfinance bank, were employed to assess the creditworthiness of a nationally representative sample in Indonesia (basing their judgments on expected returns to investments and on the stability and predictability of household cash flows). Households with incomes above the poverty line were deemed far more likely to be creditworthy than poor households. Still, the loan officers identified 38 percent of poor households as being ready and able to borrow from Bank Rakyat Indonesia with existing financial products. Johnston and Morduch (2008) conclude that the right question is not the one that has generated debate: Are the poor and very poor as a group creditworthy? Rather, the key question is: How many? And, most importantly, can the creditworthy portion be cost-effectively identified and served?

Evidence from estimating profit functions

Researchers measuring returns to capital run into the same difficulties that make impact evaluations so challenging. The biggest hurdle is to disentangle the pure return to capital (i.e., the improvement in profit that occurs relative to a situation where all else is the same, but the business owner has less capital) from the effect of qualities and conditions correlated with having capital. People with better access to capital tend also to have better access to other resources like labor and markets. They may also be more entrepreneurial, less risk averse, and higher skilled. So when we see that people with more capital have higher profits, it doesn’t necessarily mean that having more capital caused the higher profits. The gains may be due to the other attributes.

Two approaches are taken to measuring returns to capital. The first approach uses econometric methods to estimate profit or output functions, and identifies returns to capital parametrically. Identification then turns on the extent of control variables for typical confounding variables like basic ability and entrepreneurial skill. Feder et al. (1990) provide an example. Their model of farm production in Jilin province in northeast China uses a switching regression (following Maddala, 1983) in which farm households are assumed to face a binding liquidity constraint (case 1) or to be unconstrained (case 2). The two cases are determined endogenously, modeled as a probit in which the dependent variable is an indicator of credit constraints. The statistically significant variables in the probit are last season’s income and current savings levels, and, identification rests on the

---

7 See also Banerjee and Duflo (2005) for a parallel survey of the literature on returns to capital in developing economies.
authors’ assertion that neither directly affects output once capital is accounted for in the output equations. Feder et al (1990) find reason to think that liquidity constraints bind: surveys yield that 41 percent of farmers with access to formal finance indicated that they would like to borrow even more, and 28 percent of non-borrowers wished to borrow but were denied access. But the estimates yield that one additional yuan of liquidity would yield only about one quarter of one yuan of additional output. Feder et al (1990) are left to conclude that constraints may not in fact bind so strongly in practice and that a fair amount of “production credit” is likely getting diverted to consumption purposes (about a third, they estimate). Their conclusion is thus relatively pessimistic about the general proposition that financial access will raise incomes in dramatic ways (although the welfare gain from consumption smoothing should not be ignored).

Newer work is more optimistic; indeed, some of the estimated returns to capital are puzzlingly high. In a study that closely follows from Feder et al (1990), Guirkinger and Boucher (2007) use a switching regression to identify constrained and unconstrained farmers in Peru, yielding an estimate that implies that relaxing credit constraints would raise the value of output per hectare by 26 percent. The result is, of course, hypothetical, but suggests the possible gains in efficiency from expanding access to finance. In keeping with this result, Udry and Anagol (2006) also find high returns to capital in a sample of small-scale farmers in Ghana. Farmers growing non-traditional crops generated returns to capital of 250 percent per year on the median-sized plot. Farmers growing traditional crops generated returns of 50 percent per year on the median-sized plot.

In turning to small enterprise, rather than farm finance, McKenzie and Woodruff (2006), use data from the Mexican National Survey of Microenterprises (ENAMIN) and find marginal returns to capital in the range of 10-15 percent per month for the smallest firms - i.e., those with capital stocks of less than US$200. Each $100 of extra investment raises earnings by $10 to $15 per month, a handsome profit. Firms with capital stocks above $500 have a more modest average marginal return to capital of 3 percent to 5 percent. These results are robust to a wide range of controls for ability, and emerge using a semi-parametric estimator that allows substantial freedom in the estimated pattern of returns. The pattern leads McKenzie and Woodruff (2006) to reject the notion that production is characterized by important non-convexities in production here, and thus they rule out technology-based poverty traps. Instead, like the case of Mrs. Vidalia Mamami, the micro-entrepreneur described above, production can be expanded incrementally, as with her move to sell vegetables and condiments as part of her fruit-selling business. McKenzie and Woodruff also find little to suggest that the smallest businesses are particularly risky or newly established. The high returns thus seem to be bound up with capital constraints. The puzzle is that if returns are indeed so high for the poorest entrepreneurs, then why have they not saved their way out of those constraints (a point developed by Armendáriz and Morduch, 2010, section 6.4)?

One possibility is that the McKenzie-Woodruff (2006) measures are over-stated. Unmeasured ability might partly drive the results, a problem that panel data alone cannot
fully remove (since changes in capital stocks over time would likely be affected by unmeasured factors like demand shocks; McKenzie and Woodruff, 2008).

This last problem is part of a larger challenge in understanding connections between informal labor and capital markets. At a basic level, the studies here do not account for the time that small-scale entrepreneurs put into their businesses: enterprise profits are generally measured without accommodation for the value of unpaid labor, though it is often the most important input into production. Without more complete data, we cannot determine the degree to which high returns to physical capital in fact reflect returns to both physical capital and unmeasured human capital. Accounting for unpaid labor is challenging given difficulties measuring the quality of labor inputs, and a first useful step would be to put a bound on the effects by re-estimating enterprise profits under alternative assumptions about the value of own-labor. Samphantharak and Townsend (2008) offer a well-structured framework for measuring enterprise profits that draws on accounting principles used by corporations; it yields clarity, for example, on how to treat income and expenses made in different periods. Ideally the framework would be extended to fully address the cost of labor.

**Evidence from field experiments**

A second approach uses experimental methods to generate exogenous variation in capital usage. The new work attempts to address econometric problems by creating interventions that distribute capital in poor communities based in part on a randomized process. In these interventions, some people get larger transfers, some smaller, depending on a decision formula that leaves an important part of the allocation to chance. de Mel, McKenzie and Woodruff (2008b), for example, study 408 small firms in Sri Lanka and offer them a range of cash or in-kind prizes (the in-kind grants are either equipment or inventories, selected by the business owners). The prizes (worth either roughly $100 or $200) were large enough to make a difference to the businesses, all of which functioned with capital investments under about $1000. The researchers picked winners and losers using random numbers. The random element (which means that people get access to capital independent of whether they are more talented, more connected, etc.) provides a key to estimating the pure return to capital. The real returns to the capital infusions in Sri Lanka turned out to be about 60 percent per year— an impressive return, especially given that nominal interest rates on loans are 12-18 percent per year.

Experimental evidence from Mexico also shows high returns to capital, especially for smaller businesses. While the earlier ENAMIN-based study had showed returns to capital of about 15% per month, a follow-up experimental study in Leon, in the state of Guanajuato, yields average returns of 20 to 33 percent per month for small, male-owned retail businesses with no employees other than the owner (McKenzie and Woodruff, 2008). Businesses that are identified by their owners as being financially constrained, moreover, have estimated returns to capital of 70 to 79 percent per month—and these businesses are most likely to be run by poorer households. Assuming that financially
constrained households do indeed tend to be poorer, the result suggests that poorer households have a greater ability to pay for capital than better-off households, and it makes interest rates of even 10 percent per month seem reasonable in this context (though the result says nothing about female-owned businesses, nor small-scale enterprises engaged in services or manufacturing).

What do average returns tell us?

The averages, though, cannot speak to what is perhaps the biggest policy debate within the microfinance community – is microcredit an effective tool for the very poor (or should the focus be on households with incomes only slightly below poverty lines and above them)? The de Mel-McKenzie-Woodruff (2008) result yields that returns are higher for a group identified as being constrained than one that is not, but the mapping into levels of poverty has not been done. We still have little sharp evidence to adjudicate whether it is true, as Vijay Mahajan, the founder of BASIX in Hyderabad, India has said in summing up the early academic literature, that most microfinance borrowers starting below the poverty line “end up with less incremental income after getting a microloan,” and that borrowing “seems to do more harm than good to the poorest” (cited by Tripathi, 2006). Or can microcredit be a powerful tool to help the very poor, as long-argued by Muhammad Yunus and others?

One important—and surprising—result of disaggregating the evidence in Sri Lanka yields that the average impact when female-owned businesses got more capital was “not different from zero” (de Mel, et al., 2008b; 2008a). The heterogeneity is remarkable: 59 percent have returns less than zero, 14 percent in the 0 - 5 percent per month range, and only 27 percent of women had measured returns to capital over 5 percent per month. (The higher returns were for poor women with high cognitive ability, as measured by a test of number recall). Men do better on average, but about a fifth of male owners generate returns below market interest rates. The finding poses a puzzle in the context of the strong focus on serving women by many microfinance institutions (Armendáriz and Morduch, 2010, ch. 7), and it calls for further investigation in other settings.

A piece of the puzzle is found in Emran et al. (2007). They point to the role of a missing (or imperfect) labor market, especially for women, as one explanation for the success of the microfinance movement. Where the labor market opportunities for women are under-valued, it is argued, possibilities for self-employment are particularly appealing. Under these assumptions, women are willing to pay high rates of interest to microfinance institutions given limited outside options, even when their returns to capital are relatively low. Their model also shows that if using extra capital requires expanding the scale of operations, the expansion could be far more difficult for women than for men as it would require a discontinuous jump in wages if expansion requires the business to moves to hiring outside workers at market wages, becoming a net demander of labor.

Sensitivity to interest rates

18
Another vantage on the ability to pay for loans comes from a parallel set of studies on the sensitivity of loan demand to interest rates. Researchers investigate loan demand directly, rather than focusing on the nature of underlying profit functions. The existence of high returns to capital in poor communities, of course, undergirds arguments that poor households can pay high interest rates—rates that are high enough to allow microlenders to sustain themselves without donor help. A recent survey of about 350 leading microfinance institutions finds most institutions charge interest rates and fees clustered between roughly between 20 to 40 percent per year, after taking inflation into account (Cull, et al., 2009b). To some observers, these rates are very high and deserve justification. One defense has been that given high returns to capital, households are happy to pay seemingly high rates of capital as long as it is reliably delivered and in sizeable volume (Rosenberg, 2002).

Dehejia et al. (2009) investigate changes in loan demand when a microfinance institution in Dhaka raised interest rates on loans from 2 percent per month to 3 percent per month in one branch but not in a similar location. Using the comparison across branches before and after the price change, they find a substantial short-term reduction in demand (roughly a unitary elasticity), although responsiveness is moderated over the longer term. In Dhaka, prices clearly matter to customers. From an institutional viewpoint, the price hike raised revenues that allowed it to cover costs and stay afloat—and was thus seen as a clear benefit to the lender, even if some customers shifted their behaviors. The Dehejia et al. (2009) study opens an important set of questions, but relies on an assumption of comparability across branches to identify the elasticity.

Karlan and Zinman (2008a) take an experimental tack to answer the same basic question. In one case, they work with a consumer lender in South Africa who charges very high interest rates for installment credit (charging nearly 12 percent per month as interest). The researchers measured clients’ sensitivity to interest rates by mailing out over 50,000 credit offers to customers, with the letters offering interest rates that were selected at random. Borrowers turned out to be less sensitive to changes in price than expected (the elasticity is modestly negative), a finding consistent with the notion that the lenders’ customers have limited outside options for access to finance—a natural finding given the high prices routinely demanded for this lenders’ loans. While the lender is atypical (well off the charts in terms of interest rates when compared to typical microfinance interest rates), the methods and questions are essential.

In a second case, Karlan, Mullainathan and Zinman (2009b) work with a large for-profit bank in Mexico, Compartamos, and randomize interest rates at the community level in 80 geographic clusters (containing 138 branches) across the country. This allows for longer-term effects, both in terms of the decision-making process for individuals as well as competitive responses. Half of the geographic clusters were randomly assigned to receive a 0.50 percentage point reduction in their monthly interest rate (which translates to about a 10 percentage point reduction in the annual percentage rate). Customers turned out to be strongly sensitive to interest rates: the price reductions led to more clients, both new and retained, as well as larger loan sizes. The demand response was high enough to generate higher gross revenue: increased number of loans and size of
loans outweighed lost revenue from lowering the interest rate on those who would have borrowed anyhow. Costs increased with the higher volume of lending, but not as much as revenues, thus profits increased.

The question begs for replication, as the elasticity should depend on underlying fundamentals of the economy, competitive landscape, financial literacy, disclosure policies, and population. Thus as with many empirical questions, more work needs to be done in a broader range of settings.

A final caveat

There is a final caveat to bear in mind. As noted in the context of Johnston and Morduch (2008), Bank Rakyat Indonesia loan officers focus in part on whether households have the cash flow (taking into account all of a household’s income, from small business and from employment) to repay loans reliably. The loan officers are not focused exclusively on the ability to repay microcredit loans only from small business profits. Poor households have a wide variety of financial needs that go beyond enterprise—for example, financing healthcare, paying for school fees, and facilitating purchases of consumer goods. Not all poor households even run small businesses, especially in urban areas, and the poorest members of villages are often landless agricultural laborers. Thus, getting business loans is not always a priority for poor households, and the focus on business loans risks blinding policymakers and practitioners to a wider range of opportunities (Collins, et al., 2009; Johnston and Morduch, 2008).

Answering narrow questions around the sizes of returns to capital in enterprise is necessary for assessing the degree to which the available evidence supports basic claims—namely that when borrowers do fund small businesses, the profits are sufficient to justify the interest rates charged by microlenders. However, given the plethora of non-enterprise needs for access to credit, it is worth noting that generating returns to capital in enterprise is only one part of what determines the demand for loans.
5. Credit market innovations

Theorists and practitioners often assume that problems, such as information asymmetries and difficulties enforcing legal contracts, undermine credit markets in developing countries. Yet in practice, deep methodological challenges have frustrated social scientists attempting to study the existence, extent, and nature of such problems. This is not to say there is evidence against the existence of market failures. But the mere fact that a micro-entrepreneur does not use credit is not sufficient cause to argue that a credit market failure exists. Several steps must be shown empirically in order to identify a market failure, form well-reasoned policy prescriptions, and finally to assess the welfare implications of resolving the market failure.

The simplest evidence of a credit market failure comes from evaluation of interventions which show increased use of credit after some shift in supply. This has been shown both for microenterprises (Banerjee, et al., 2009 in India, and; Karlan and Zinman, 2009a in the Philippines), consumers (Karlan and Zinman, 2009b in South Africa) and small and medium enterprises (Banerjee and Duflo, 2008 in India). Had these studies showed instead a substitution from one source of credit (presumably more expensive on some margin) to another, the evidence of credit constraints per se would not have been as clear.

However, noting the mere presence of credit constraints is not sufficient for policy. The following section attempts to illuminate four essential questions regarding a presumed credit market failure: (1) What is the exact nature of the presumed market failure? (2) Is there a particular policy prescription that would mitigate a particular type of information asymmetry? (3) Does this particular policy solve, or diminish, the market failure? And (4) does it lead to a welfare improvement?

The previous section addressed the question of whether or not the lack of credit we observe in developing countries is economically efficient. Specifically we explored whether entrepreneurs are borrowing less than optimally because of a lack of profitable opportunities or whether, instead, there are other constraints at play which limit demand for credit. We presented evidence from experiments in developing country settings to suggest that returns to capital are in fact often high on average, but that heterogeneity in returns, due to different levels of social connections and human capital, may lead to low demand by some for entrepreneurial credit. An important area for further research is to understand the nature of heterogeneity in returns and the relevance of such heterogeneity for assessing the extent of market failure.

The interaction between informal and formal markets will recur as a theme in the discussion below. Merely observing a formal market failure need not lead to inefficiency if the informal market is complete. Examining the mechanics of the informal market is crucial for two reasons. First, the strength of the informal market is important for measuring and predicting how effective specific formal sector interventions will be. Second, lessons learned in the informal markets can help shape policy in the formal
markets. Group lending, for example, is based largely on lessons learned from observing risk-sharing and credit and savings associations in informal markets.

A second important theme is the application of the theoretical literature on information asymmetries to consumption loan markets. While theories on information asymmetries are often applied to discussions of entrepreneurial credit markets, the same theories are relevant to consumer credit as well for several reasons. As a start, the line between entrepreneurial “investment” and consumption “smoothing” is rarely evident for small, closely held businesses. Money is fungible. Empirical evidence from Johnston and Morduch (2008, Table 6) shows how even entrepreneurs report using loan proceeds to pay for consumption. In data from Indonesia, clients of BRI reported that about half of their loans were being used for non-entrepreneurial activities. More generally, asymmetric information problems as applied to risky “projects” have natural and close analogs for consumption loan borrowers. Just as entrepreneurs may respond to interest rates according to unobservable fixed characteristics of the return structure of their “project”, so too may the elasticity of demand depend on unobservable fixed personality traits (e.g., trustworthiness) or behaviors (e.g., the probability of incurring bad shocks). These unobservables may have significant impact on the likelihood of repayment. Similarly, if entrepreneurs unobservably change effort levels or repayment choices in response to interest rates, individuals may change their effort in activities such as maintaining wage employment or securing alternative sources of cash in the event of a bad shock. And, of course, individuals may also default strategically.

5.1 Nature of Frictions and Policy Examples

Understanding how and why markets fail to provide credit access for the poor is critical to designing and replicating innovations that improve access. For example, if the problem is adverse selection, then policies that help reveal hidden information should help lenders screen and tie prices appropriately to risks. Such policies can be at the product or process level within the firm, or at the regulatory level, with credit bureaus. For example, Ghosh and Ray (2001) and Drugov and Macchiavello (2008) show theoretically how small, initial “tester” loans can provide information to lenders useful for assessing risk on subsequent, larger loans. Karlan et al. (2008a) show in a field experiment how social networks can be used to generate referrals of “good” clients, improving loan repayment.

Identifying the specific nature of a market failure is difficult, however. Take, as an example, information asymmetries generated through the interest rates on credit. Even if a lender were to randomize its interest rates across borrowers, merely observing that high rates lead to higher default does not disentangle selection from moral hazard and income effects. In a study of consumer loans in South Africa, individuals were marketed a randomly assigned interest rate and then, following the loan take-up decision, the lender randomly assigned a contract interest rate (lower than the initially offered rate), and a dynamic incentive (Karlan and Zinman, 2009b). This double-tiered randomization (before and after selection into credit) provided for a separation of adverse selection from
moral hazard, and perhaps most importantly, provided a roadmap for using experimentation to test theories in which selection is useful to separate from ex-post incentives. Karlan and Zinman found weak evidence for adverse selection, only statistically significant for females, and stronger evidence for moral hazard, in particular the dynamic incentive (lower interest rate on future loans).

In some cases, the hidden information is not hidden per se, but rather just not used. Interventions that work with lenders to improve their data processing can also improve efficiency. This has been shown in both insurance markets (Finkelstein and Poterba, 2002) and credit markets (Karlan and Zinman, 2009b). Credit bureaus also provide institutional and market level mechanisms for revealing hidden information. Theoretically, it is also the case that introduction of credit bureaus can mitigate moral hazard, as the consequences of default increase if lenders share information with each other (de Janvry, et al., 2008). On the other hand, public credit bureaus can have adverse effects, leading to coordination failures (Hertzberg, et al., 2008). We will discuss these more below.

If the problem is moral hazard, on the other hand, solutions may differ. Again, one could categorize solutions at the business or product level, and separately at the institutional or market level. At the firm level, policies include contracts with dynamic incentives, tighter enforcement, or even business or financial training to improve the outcomes of the micro-enterprises or cash management of the borrowers. Note that ex-ante moral hazard, i.e., choice in effort, is often construed as an analytically derived, forward-thinking “decision” by the agent. A set of behavioral stories also fit into the ex-ante moral hazard category, albeit without a “conscious” decision to shirk. For example, someone who succumbs to temptation (hyperbolic, quasi-hyperbolic) and thus finds herself without sufficient cash to repay a loan would fall under ex-ante moral hazard. Thus interventions aimed at addressing cash management of borrowers should be categorized as interventions to address ex-ante moral hazard. At the institutional level, interventions include improving legal institutions for enforcement as well as improving internal banking information and payment systems so that banks can engage in enforcement activities more efficiently.

We now review the three main categories of market failures. We note however that although the theory as currently written clearly separates these mechanisms, in practice markets may, and likely do, contain elements of all three. For example, high interest rates may attract individuals who intend to exert less effort (as compared to the traditional adverse selection story that high interest rates attract individuals with risky projects). We will discuss such interactions below in more detail.

**Adverse and Advantageous Selection**

---

8 “Adverse selection” in this context also includes “lower anticipated effort” which may be generated through the same pooling process that traditional adverse selection on risk-type generates. Such an “anticipated effort” effect requires moral hazard to be possible, since if effort is perfectly observable this effect would not occur. However, hidden information must also be present for this mechanism to be relevant, since otherwise firms would price according to anticipated effort by the borrower.
Lenders set interest rates to maximize returns and may charge higher rates to riskier borrowers to compensate for higher anticipated default. Stiglitz and Weiss (1981) demonstrates that under imperfect information (in particular about borrowers’ production functions) higher interest rates can lead to adverse selection, as only borrowers with riskier projects (and higher returns in a positive state) will be willing to pay such prices. Moreover, by lowering returns in all states, higher interest rates can induce all borrowers to undertake riskier activities.

Advantageous selection is also a possibility, and if such models accurately describe credit markets, would lead to very different policy implications. de Meza and Webb (1987; 1989) discuss the mechanics of such a model, and the difference between their prediction and that of Stiglitz and Weiss (1981). The key difference lies in the assumption about the relationship between risk and return. The de Meza and Webb papers assume that entrepreneurs with higher intrinsic quality have higher returns that first-order stochastically dominate lower-quality entrepreneurs. This leads to the opposite result of adverse selection: as a bank raises its interest rate, the marginal client that drops out is a low-quality client.

Boucher and Carter (2005; 2008) develop a model demonstrating that imperfect information can lead to a form of rationing of credit in which would-be borrowers are deterred based on the terms of the contract, rather than price. Under this model of “risk rationing,” by raising collateral requirements lenders shift so much contractual risk to borrowers that even borrowers with the collateral required to qualify for a loan choose not to borrow under these terms for fear of losing vital collateral. Note that “risk” to the borrower is both about explicit risks such as loss of collateral, whether physical or social, but also is about “ambiguity,” where “ambiguity” describes the borrower’s fear of dealing with formal institutions with which borrowers have little experience. Ambiguity here may also describe the fear of being stigmatized by formal institutions for being poor. Breaking through these barriers may require significant bravery on the part of borrowers. This model assumes that the risk-averse clients, the ones who refuse to borrow at offered terms because they fear the mental anguish of default or the risks of unknown or ambiguous legal and social consequences, are also the low risk clients in terms of the projects they pursue. As a result, high-return entrepreneurship is limited, particularly among the less wealthy, and welfare is reduced, as risk-averse agents choose safer but lower-paid employment or lower-return entrepreneurial activities.

**Ex-ante moral hazard (effort)**

A second source of friction in credit markets is ex-ante moral hazard, or effort. This can take on many forms. By lowering the net profits for a given business, higher interest rates, combined with limited liability, reduce the incentive to invest extra effort in production. Reduced effort diminishes returns to borrowers even when projects are successful and also increases the probability of failed projects and, by reducing the probability of sufficient returns, increases the likelihood of default.
Note that increased default at higher rates is not necessarily due to traditional entrepreneurial shirking or to adverse selection as discussed above. While the lender sees only whether or not the loan is repaid, there may be more nuanced explanations for defaulters’ lack of sufficient cash flow to make payments. In household finance, higher interest rates may affect borrowers’ effort to retain or obtain employment, to tap alternative sources of cash (e.g., borrow from a family member) in the event of a bad shock, or to manage consumption in order to retain sufficient funds for loan repayment. These examples do not relate to a firm’s production function, but the same logic applies: at higher interest rates, consumers may exert lesser effort to make sure they have the cash available to repay the loan.

The above examples all pertain to conscious shirking. However, other mechanisms may be at play, which exhibit the same reduced form prediction: insufficient cash is available at the time of repayment, and the lack of cash is a result of decisions of the borrower. For instance, temptation models predict that, under uncertainty, some types of consumers will consume more in the future than they consider optimal at the time of their decision to borrow. Thus they will have less money available to repay the loan.

Lack of entrepreneurial skills is another source of information asymmetry. If entrepreneurial skills are fixed but unobservable characteristics, one could argue this is also appropriately labeled adverse selection since the lender cannot tell ex-ante which individuals will likely put in effective effort. The problem is that the lender here cannot enforce “effort” (e.g., being creative, being assertive), similarly to the discussion above regarding Karlan and Zinman (2009b) with respect to interest rates and information asymmetries.

One way to generate insight into whether entrepreneurial skills are fixed is to try to teach these skills. Karlan and Valdivia (2009) conducted a field experiment in which microcredit borrowers were randomly selected to receive free business training on topics including cash management, business selection, and marketing. Results were mixed. Treated individuals earned greater profits in their businesses, and in particular greater profits in bad months, although this result was not robust to alternative econometric specifications. Along with marginally improved repayment, client retention increased significantly, suggesting the clients placed great value on the training they received. These institutional benefits led training to be profitable for the microfinance institution, even with no cost-sharing on the part of clients.

*Ex-post moral hazard (voluntary default)*

Voluntary default is distinct from ex-ante moral hazard because here the borrower has the ability to repay and chooses not to do so. Thus the problem lies strictly with enforcement of repayment. In the simplest construction, with respect to interest rates, borrowers are more likely to default voluntarily at higher rates than at lower rates since the cost of repaying rises, but the benefits of repaying remain the same. Much of the rhetoric around lending to the poor, including Yunus’s strategy with respect to the Grameen Bank,
assumes that voluntary default is extremely rare and that entrepreneurial individuals will repay as long as they can.

Demonstration of the presence or absence of ex-post moral hazard requires identifying situations in which, despite the availability of cash, individuals default on loans by choice. Provided good survey measures, observational data on borrowers should allow us to identify whether default occurs even among clients who have the funds to repay. One could also identify ex-post moral hazard indirectly, by showing that “trustworthy” individuals are more likely to repay loans. In Karlan (2005), individual borrowers from FINCA, a Peruvian micro-lending program, played a “trust” game and completed the General Social Survey, which contains questions about trust, fairness and helping others. Both the survey and the “trust” game were intended to measure individuals’ trustworthiness. The “trust” game was conducted with two players and an administrator. First, each player received three Peruvian soles coins. Player A is then instructed pass to Player B any or all of her three coins, which the administrator then doubles. Player B then chooses how much money to return to Player A in response (the administrator does not double the return). Individuals who return more when playing as Player B than when playing as Player A are considered trustworthy. Repayment records from the entrepreneurs’ microfinance association show that individuals who show trustworthiness in this game exhibit lower default, lower drop-out, and higher savings balances. The survey yielded similar results, showing that individuals who respond more positively to General Social Survey questions about trust, fairness and helping others, are more likely to repay their loans one year later.

Drawing such inferences about personality types can lead to a gray area between moral hazard and adverse selection that the theoretical literature has yet to delineate. In the classic case of adverse selection individuals with riskier production functions are more likely to borrow at higher rates, and less likely to repay. The effort borrowers invest in ensuring the success of their projects might normally be associated with moral hazard, but if higher interest rates attract borrowers who are inherently apt to expend less effort in their businesses the issue is really selection. Moral hazard is here interlinked with adverse selection.

5.2 Interventions and mechanisms

*Group lending in theory*

Group lending is perhaps the first and most oft-discussed “solution” to information asymmetries in developing countries. Adverse selection and moral hazard are dealt with by effectively shifting the responsibility of screening, monitoring, and enforcement from the lender to clients. Group liability requires that if one group member defaults, her fellow group members will be responsible for her payment. Under group liability schemes then, clients have an incentive to screen other clients so that only trustworthy individuals are allowed into the program. Ghatak (1999) describes how group lending can, in theory, mitigate adverse selection through group formation. Potential clients, selecting fellow borrowers with whom they will be jointly liable for loans, will exploit
information known to borrowers but not to banks to screen out bad credit risks. Groups segregate according to riskiness, with riskier borrowers joining with other risky borrowers. Thus, the cross-subsidization of risky customers by safer customers that is presumed to be inherent in individual liability schemes is mitigated. Group lending also addresses moral hazard by providing incentives for clients to employ peer pressure to ensure that funds are invested properly and effort exerted until the loans are repaid in full. By lowering default, the expected total cost of borrowing for borrowers can be reduced, improving welfare especially for households without collateral. Still, because clients face the potential of default from fellow group members, group liability could also be seen as a tax, effectively increasing the net interest rate on safer borrowers (Stiglitz, 1990).

Just as social networks can facilitate informal insurance mechanisms between households, borrowers, who often know each other, can serve as mutual insurance pools for fellow borrowers, covering loan payments when a fellow group member is unable to. Moreover, their familiarity should allow for mutual monitoring to avoid making payouts due to moral hazard. If social networks are important in fostering repayment under group liability we should see higher repayment where borrowers are more connected to each other. Wydick (1999) provided early evidence of this in Guatemala, testing the effect of three different types of social cohesion: peer monitoring, social ties, and borrowing group pressure. Wydick finds monitoring to be the most important of the different forms of cohesion, with little effect from social ties. Ahlin and Townsend (2007), with data from Thailand, find social ties (measured by sharing among non-relatives, cooperation, clustering of relatives, and village-run savings and loan institutions) actually reduce repayment rates, though stronger social sanctions improve them.

The conclusions one can be safely draw from these findings are limited however; because borrowers self-select into groups group formation is determined by factors endogenous to the forces of liability scheme itself. If, for instance, borrowers with successful enterprises have greater social networks there may be a correlation between social ties and repayment that has nothing to do with group monitoring, enforcement, or cohesion, but rather is due to success in entrepreneurship. Karlan (2007) uses quasi-random variation in the group-formation process at a Peruvian microfinance institution to show that groups with greater levels of social connection (ethnic ties and geographic proximity) have lower default and higher savings rates. Greater knowledge of default status among fellow borrowers, and deterioration of relationships after default, suggests both monitoring and enforcement play a role.

The structure of group-liability contracts produces an ambiguous effect on risk-taking in entrepreneurial investments, and thus on returns (Fischer, 2008). The insurance aspect of joint liability creates an incentive to take risks by lowering the cost of default, while the monitoring and enforcement (group pressure) aspects increase the cost of default. To generate empirical data on the balance of the effect, Fischer ran a series of investment games in which microfinance clients in India “borrowed” and “invested” according to different types of contracts where players could share risk by making income transfers to their partners. Returns were randomized and payouts from the contracts were in real money of up to a typical week’s income. Each treatment was conducted under both
complete information, where all actions and outcomes were observable, and limited information, where players observed only whether their partners earned sufficient income to repay their loans.

Several key implications of joint liability emerged from the results. First, joint liability produced free-riding: risk-tolerant individuals (as measured in a separate risk game) made significantly riskier investments under limited information, without compensating their partners for the insurance they provided. However, under complete information joint liability did not encourage greater risk-taking. Peer monitoring via approval over partners’ investment decisions mitigated *ex ante* moral hazard by discouraging risky investment choices.

Fischer’s results are broadly consistent with those in Giné et al. (2009), generated from games conducted with entrepreneurs in Lima, Peru. Here, too, joint liability encourages free riding, but the mutual insurance it provides prevents default from increasing. Moral hazard is reduced by allowing clients to form their own two-person groups, but interestingly there is no apparent pattern of matching between risk-averse partners. Instead the effect seems to be one of fairness: though safe-risky investment pairings are viable over the long term, and generate higher returns, both partners opt to either both choose risky investments or both choose safe investments. This leads to suboptimal risk-taking overall. The policy conclusions from these findings depend on the returns to real-world investment choices, but the assumption that riskier projects would be rewarded seems reasonable.

Conning (2005) develops a model to compare outcomes under both individual and joint liability and finds no clear winner: each is optimal under different circumstances, depending on the type of borrower. Interestingly his conclusions do not rely on an information advantage of borrowers over bank delegates. The microfinance industry collects a great deal of information on the financial performance of its institutions, and this data is helpfully broken down by institution type, including lending methodology. Drawing useful inferences from the cross-sectional data about which specific interventions, like group lending, mitigate information asymmetries can be difficult because of the difference in target markets served by each lender type. Individual loans, for instance, are most often offered to borrowers at the upper tier of the microfinance market. These borrowers are likely to have levels of assets (including working capital, human capital, and collateral) and production functions distinct from the type of entrepreneurs typically served by group-lending institutions. Interest rates, too, are correlated with poverty levels: the fixed costs of processing loans imply that interest rates will be higher on smaller loans.

Cull et al. (2007) use a dataset comprising 124 microfinance institutions across 49 countries to analyze MFI performance and outreach. They find patterns in the data in line with the theoretical literature on adverse selection and moral hazard, but only for certain types of institutions. Specifically, lenders making individual loans show higher portfolio-at-risk (PAR) as interest rates increase, and after a certain point (60 percent annualized percentage rate) profits fall off. Reduced demand at higher rates likely contributes to this
effect. This trend does not hold however for group lenders, suggesting that they are better able to mitigate problems stemming from information asymmetries.

**Group lending and related mechanisms in practice**

While lenders have been attracted to the intuitive appeal of relying on borrowers to monitor themselves, through group enforcement, recent successes of several MFIs offering collateral-free individual-liability loans to poor borrowers have caught the attention of the microfinance industry. While individual-liability schemes cannot overcome the problem of information asymmetries between client and lender, the methodology does avoid other problems typical of group lending, such as strategic default among group members. Nevertheless, the strong financial performance of a few lenders is not sufficient evidence to advocate the end of group lending. The low default rates among existing individual-liability borrowers may be indicative of superior management skills among these few programs, heroic (or coercive) efforts by field staff to ensure repayment, a culture of repayment in these regions (notably Bangladesh), or credit constraints among clients (who want to maintain access to future loans).

Armendáriz and Morduch (2000) report on microfinance institutions in Eastern Europe and Asia that use a combination of mechanisms, such as rebates, larger loan sizes, and faster loan approval for safe clients, making individual lending possible and profitable. Whether the bank employs incentives or sanctions, in theory the bank will extract payoffs up to the borrower’s opportunity cost of not repaying. The size of the opportunity cost will depend on borrowers’ outside options for credit. It can therefore be difficult to predict borrower behavior under different contracts. In two experiment with a microlender in the Philippines, Giné and Karlan (2009) first randomly assign groups of joint liability clients to either remain under joint liability or to have their loans converted to individual liability, and second randomized before the initial creation of lending groups whether they would be group or individual liability lending groups. All other aspects of the loan contract—interest rate, payments, and term—remained the same. For the first experiment, follow-up of the study participants after three years revealed that converting from group to individual liability had no adverse affect on default rates, despite the fact that the conversion reduced peer monitoring of loans (note that group meetings remained intact, just the group liability was removed). Furthermore, results showed greater client retention among individual-liability borrowers, making the switch appear profitable for the bank and appealing to clients. The second experiment then also allowed for selection effects from the group liability, by pre-announcing before the creation of groups whether or not the village was eligible for group or individual liability loans (again, keeping group meetings intact, and just varying the liability rules). Similar to the first experiment, repayment rates remained high, near 100%. However the credit officers in this second were less likely to enter certain villages and create groups, thus indicating fear of default from the bankers' perspective. This could either be because credit officers know the relative risks, or due to the lack of learning by the credit officers that the group liability is in fact not a necessary component to generate high repayment.
Rutherford (2000a) argues that “financial services for poor people are largely a matter of mechanisms that allow them to convert a series of savings into usefully large lump sums.” The difference between credit and savings is simply whether the lump sum is taken at the beginning or the end of the cycle. The deposits are small, of necessity, while the lump sums are used for all sorts of expenditures: life-cycle events (births, school fees, home-building, deaths), emergencies, and investments. By way of describing several types of financial institutions for the poor, from a homegrown “merry-go-round” savings scheme to an NGO-created MFI, Rutherford maps out the relationship between the complexity of the institution and the flexibility of the products they offer. The more formal the institution the more likely it is to be able to turn savings into a lump sum at a time that is convenient for clients. With simple savings clubs members must wait for their turn at the lump sum handed out each week, and often the club must be reformed at the end of every cycle. But there can be tradeoffs on the other end too; the local moneylenders and deposit collectors of the informal market provide services with convenience, a measure on which formal sector services have fallen short.

Seen in this way, credit can be a useful mechanism to help people save toward the purchase of an asset. Microfinance programs emphasize small, frequent, regular payments, and create incentives for clients to make those payments. For those who have trouble making regular savings deposits, borrowing can be a way to commit to making those payments (albeit at greater total cost to the borrower). But are bank contracts sufficient, or do these savers need the social pressure of group loans to continue the payments? Basu (2008) develops a model showing hyperbolic discounters, who place a greater value on present consumption than future consumption (and therefore find it harder to set aside savings) will prefer to remain in ROSCAs even when there are no sanctions for absconding with their loans—because it is preferable to have the commitment device of the ROSCA over the long run than free money today. Following a similar logic, Bauer, Chytilova, and Morduch (2008) use data from south India to show that hyperbolic discounters borrow more and save less than others, conditional on household characteristics, and that they are especially likely to borrow from microcredit programs which offer helpful structure and social support.

Rutherford emphasizes flexibility in addition to stability and security as important factors in clients’ take-up decisions, but it is unclear which institution types are best in this regard. While some MFIs offer a range of loan, savings, and insurance products, others offer only one: for example, a four- or twelve-month enterprise loan with a weekly payment schedule that may not match the cash flow of the poor. It is striking how many MFIs require loan repayments beginning only two weeks after disbursement—and how few microfinance account officers believe their clients’ investments will start to pay off by that time (Armendariz de Aghion and Morduch, 2000). Moreover, there is a substantial transaction cost to conducting these meetings, for both the clients and the MFI. However, many MFI managers consider weekly payments essential for several reasons: 1) the smaller repayment amounts are easier for clients to manage; 2) it creates a culture or discipline of repayment for those not used to formal payment deadlines; and 3)
the regular meetings prevent attention problems (e.g. forgetting when the repayment meeting is supposed to take place). In a theoretical paper, Jain and Mansuri (2003) give another possible reason: that the frequent repayment schedules force borrowers to turn in part to the informal sector, which is beneficial if the banks can then piggy-back on monitoring by moneylenders.

Empirically, the question of what is the optimal repayment frequency for borrower and lender remains unanswered. Field and Pande (2008) conducted a simple field experiment in India in which they randomly assigned new borrowers either to traditional weekly payments or to monthly payments. They find no difference in repayment, although the results are preliminary (after one year), pertain to small-sized loans only, and hold for new borrowers, not long-term borrowers. McIntosh (2008) extends the Field and Pande result with a study of a Ugandan MFI in which the bank offered its village banks a choice between weekly or bi-weekly repayment. The choice had to be made by unanimous vote of each village bank’s members. A simple comparison of the repayment performance of weekly to bi-weekly banks would be laden with two sources of bias: geographic selection by the MFI, and self-selection by the clients. McIntosh constructs a comparison group by asking clients in ineligible regions to decide whether they would want to switch from bi-weekly repayment, had it been offered to them. By making two comparisons: groups that switch to bi-weekly payments vs. ineligible groups, and would-be switchers vs. those who would not choose to switch McIntosh is able to eliminate some of the bias in the estimate. He finds no drop in repayment (actually a slight improvement) and a large increase in client retention (dropout is reduced by 40%).

**Credit scoring and credit bureaus**

Adverse selection problems can be mitigated by lenders through better screening processes, a process which lenders have been continually refining. Credit scoring has become more sophisticated over time, but most developing countries still lack credit bureaus (and in many cases, unique identification numbers for citizens). It is often difficult to evaluate implementations of national programs but de Janvry, McIntosh and Sadoulet (2008) exploit a natural experiment in which a large MFI in Guatemala installed a new credit reporting hardware in waves, allowing the researchers to compare early adopting branches to late adopters. Findings from the study show that with the new technology the lender screens out substantially more applicants, but also makes more new loans. New borrowers have higher repayment and take larger loans. In related research the same authors worked with an MFI to randomly assign some clients to receive training on the importance of credit bureaus to borrowers’ credit opportunities. The clients were informed both that late payment with one lender will harm their access to credit at other lenders and that paying on time gives them greater access to credit at potentially lower rates. The authors find that the training led to higher repayment rates by their clients, but also led their clients to borrow elsewhere after establishing a good credit record.

While credit bureaus can clearly pay a valuable role in leveling information asymmetries, they nevertheless are only capable of providing data. Lenders must still interpret the data and decide whether to approve loans, and at what terms. Working with a consumer-credit
lender in South Africa, Karlan and Zinman (2009b; 2009a) show that there is room for improvement in the screening process. By extending loans to randomly selected applicants who would have been rejected under the bank’s own scoring system they show an increase in outreach and profit for the lender, despite the fact that the broader client base carries with it higher default. By comparing welfare of these randomly approved marginal clients to a control group of identically marginal applicants who remain denied, the researchers find that the loans are significantly welfare enhancing for borrower households.

Dynamic incentives

Dynamic incentives have also been credited with helping to solve moral hazard problems. Theoretical work beginning with Bolton and Scharfstein (1990) has shown the conditions under which dynamic incentives, for example the threat of termination of credit, can generate sufficient incentives for the borrower to repay loans. Whereas this work established a clear theoretical understanding of the potential role for dynamic incentives, in practice specific methods of implementing dynamic incentives may or may not work, and evidence is needed to understand how this theory can be employed to improve product design. For example, progressive lending, in which initial loan sizes are capped low and progressively larger loans are only available after successful repayment, has become the norm for many microfinance lenders.

Clear evidence on dynamic incentive components of the contract structure are more difficult to show, as most lenders only vary contract terms endogenously, for selected clients. In one study in South Africa, discussed above as well, Karlan and Zinman (2009b) worked with a lender who randomly offered some clients a dynamic incentive, a discount on future loans assuming successful repayment of their current loans. This offer led to a 10% reduction in default (from a base default rate of 15%), and the responsiveness was proportional to the size of the incentive. This simple experiment naturally opens the door to further questions, such as whether the framing of the incentive matters (e.g. loss versus gain framing), and whether a dynamic incentive can continue to work over the long term or will instead collapse over time, either due to capacity constraints of the lender or to changes in responsiveness to incentives over time by the borrowers.

Controlled laboratory experiments can shed insight into questions that may be difficult to answer through a natural field experiment (for a taxonomy of field experiment methodologies, see Harrison and List, 2004). Lenders are often resistant to making the changes to operations required of a field experiment. Few lenders will, for example, agree to a field experiment in which further loans are made after default, or the reverse, to pledging not to lend any more even to good borrowers. For such tests, we turn to controlled laboratory experiments. Several papers have examined the power of dynamic incentives, on topics far afield from microfinance. Relevant papers include Charness and Genicot (2007), and, as described above, on microfinance by Giné et al.(2009) and Fischer (2008). In each of these papers, the dynamic incentive led to the largest effects on reducing moral hazard problems.
As a general point, dynamic incentives on lending contracts are not much different than conditional transfer programs. Like conditional transfer programs, dynamic incentives provide a future incentive (e.g. larger or cheaper loans rather than cash) in exchange for a change in behavior now (e.g. lower moral hazard rather than higher school attendance). The interesting questions here are not whether people respond to incentives, but rather how to structure the incentives to be socially optimal, how to frame the incentives so as to maximize their immediate and long-term effectiveness, and how to structure and time the incentives so as to maximize social welfare. On the last point, an excellent example comes from Colombia where the government tested the importance of timing in a conditional cash transfer program by randomly assigning some individuals to receive their conditional cash transfers at the time school fees were due rather than before. This is a classic example of Thaler and Sunstein (2006; 2008) employed in developing countries and provides a clear example of how the structure (in this case, the timing) of the incentive, not just the incentive itself, can affect social outcomes.

5.3 Impacts from solving credit market failures

For a poverty intervention as widespread as microfinance, with an estimated 154 million clients worldwide (Daley-Harris, 2009) and over five billion dollars invested each year (Forster and Reillie, 2008), there is surprisingly little rigorous evidence of the impact of microfinance on household welfare. In part this is due to the difficulties of measuring impact. There ample is reason to think that selection bias may be especially problematic with non-experimental evaluations of microfinance programs. Microfinance clients are likely to possess a special determination and ability to improve their welfare and therefore comparing their outcomes to the outcomes of non-clients (presumably without this drive) will overstate the impact of microfinance. Unfortunately, personal characteristics like “entrepreneurial ability” or “drive” are either difficult to measure or unobservable.

A related challenge is bias from program placement, in which outcomes in program villages are compared to outcomes in non-program villages. The problem with this method is that programs choose where they operate for a reason. They may target the poorest villages, for instance, or they may start cautiously with better-off clients before expanding their outreach. The bias from non-random program placement, therefore, can go either way, depending on whether the evaluation compares program villages to non-program villages that may be (even unobservably) better or worse off.

It will not be clear how serious a problem these omitted variables are until rigorous impact assessments of credit are completed, side-by-side with alternative assessments, and the results can be compared. As a first step Coleman (1999) shows how important selection bias can be in a study of microfinance borrowers in northern Thailand. By forming a group of prospective microfinance clients who signed up a year in advance to participate with two village banks, Coleman was able to create a comparison group mostly free of selection bias, as both the borrowers and the non-borrowers had selected into the program at the same point in time. Coleman then generates two estimates of the
impact of the program: an unbiased estimate using the clients who signed up in advance as the comparison group; and a “naïve” estimate using a group of seemingly similar non-participants (as in typical non-rigorous evaluations). Comparing his unbiased impact estimate to the estimate he would have calculated had he naïvely compared program participants to a group of non-participants Coleman finds the “naïve” estimate substantially overstated the gains from participation on several outcomes (especially women’s landholding) because participants turned out to be initially wealthier than non-participants.

In an ambitious early attempt to solve these identification problems, Pitt and Khandker (1998) surveyed 1,798 member and non-member households of three Bangladeshi MFIs (Grameen Bank, BRAC, and RD-12), and exploit eligibility criteria (landholding totaling less than one-half acre) to obtain measures of impact of credit. While there should be no discontinuity in income between people who own just over or just under a half acre of land, participation in the MFIs would be discontinuous because those who were above the cutoff would be rejected from the programs. They find huge impacts: every 100 taka lent to a female borrower increased household consumption by 18 taka. However, in a re-analysis of the data Morduch (1998) challenges the econometric models and identification assumptions in Pitt and Khandker (1998). Using a difference-in-difference model, he finds little evidence for increased consumption but does find reduction in the variance in consumption across seasons. He argues that in the cross-sectional set-up, non-random program placement is only addressed through restrictive assumptions and that reliable inferences require additional data.

Khandker (2005) refined the earlier model with the benefit of panel data, finding lower impact estimates but greater total impact (from current and past borrowing in the survey rounds conducted in 1991-2 and 1998-9) and substantially lower marginal impact from new borrowing. Poorer clients are found to have larger impacts than the less-poor, and money lent to men is not shown to have any impact at all.

Roodman and Morduch (2009) attempt to find closure to the issue by returning to the data and rebuilding the analysis from scratch. They are unable to replicate results from Pitt and Khandker (1998) or Khandker (2005). In fact, their estimates carry the opposite sign. Rather than concluding that microcredit harms borrowers, however, they unearth a raft of identification issues which are not solved with panel data. Their revised analysis casts doubt on all of the findings from the related set of papers, including Morduch (1998)’s finding on consumption smoothing.

Karlan and Zinman (2008b) conducted a randomized controlled trial to measure the welfare effects on borrowers at the same time that they worked with the South African consumer-credit lender to deepen outreach. They find the loans are substantially welfare enhancing for the borrowers, who are 11% points more likely to remain employed, are less likely to encounter periods of hunger or to score as poor, and had higher income and improved credit reports. It should be noted this is the impact on a particular set of borrowers: those who would have been rejected under the lender’s standard approval requirements.
Burgess and Pande (2005) use the introduction and eventual repeal of an Indian social banking law to evaluate the impact of access to finance on the rural poor. Between 1977 and 1990 Indian banks wishing to open new branches in locations already served by any commercial bank were required to open four new branches in un-served locations. Since the policy required banks to focus on locations with the lowest level of financial access the authors use the pre-1977 levels of financial intermediation as an instrument for the states that will be most affected by the policy. They find a one percentage point increase in the share of credit disbursed by rural branches reduces rural poverty by 1.5 percentage points while a one percentage point increase in the share of savings held by rural banks reduces poverty by 2.2 percentage points. The opening of one bank branch per 100,000 people in a rural unbanked location reduces rural poverty by 4.7 percent. As optimistic as these results are, the authors caution that default rates during the expansion period were in the 40 percent rage, leading to the abandonment of the program. As always, appropriate mechanisms must be designed for credit to be sustainable. The authors’ caution also creates uncertainty about whether the results are due to access to finance per se or from the fact that a large share of customers received money as loans but never repaid them (creating implicit cash transfers).

Several randomized studies are being conducted on both individual lending as well as randomized program placement designs, similar to the Progresa evaluation in Mexico. These differ from the above mentioned South Africa experiment in that they are traditional “microfinance” credit programs that target entrepreneurs (rather than customers looking for consumer loans). Results from the first two trials, from urban India and the Philippines, challenge popularly held assumptions about the use of credit. Other studies underway in rural India, Morocco, rural Peru, urban and rural Mexico, and China will help paint a more complete picture whether or not microfinance is welfare improving, by what measures, for whom and under what conditions.

Karlan and Zinman (2009a) uses a credit-scoring methodology to evaluate the impact of loans to microentrepreneurs in urban Philippines. The methodology used is similar to Karlan and Zinman (2008), above, but there the focus was on loans made to employees. Here the effects are much more muted, and some findings cast doubt on the traditional microfinance narrative. Business owners’ profits increase, but not through investment in productive assets or working capital. Moreover, the treatment effects are stronger for groups that are not typically targeted by microlenders: male and higher-income entrepreneurs. There is evidence that treated businesses actually shrink in size and scope, including the shedding of paid employees. The results suggest that borrowers used credit to re-optimize business investment in a way that produced smaller, lower-cost, and more profitable businesses. The question remains as to how credit enabled this change: why did business owners need to borrow to reduce staff? One potential explanation is household risk management: treated individuals substitute out of formal insurance products, while also reporting a greater ability to borrow from friends or family in an emergency. It is possible that before credit entrepreneurs were retaining unproductive employees as a kind of informal mutual benefit scheme. Those employees, even if unprofitable, were an additional place to turn in times of need.
In urban India, Banerjee et al. (2009) evaluate the impact of a non-profit group-lending microfinance program in the slums of Hyderabad. The researchers randomly assigned 120 slums to either treatment or control. After 15-18 months the households from the treatment slums were compared to the households in the control slums. The results show impacts on a number of dimensions, though not, critically, on average consumption. The treatment slums have greater investment in business durables, increases in the number of businesses started, and in the profitability of existing businesses. Among households that did not have existing businesses at the start of the program, those with high propensity to become entrepreneurs see a decrease in consumption, while those with low propensity to become entrepreneurs increase consumption. Likely this difference is explained by investment in durable goods among those likely to become business owners. While the short-term impacts are clear, this result makes it difficult to anticipate the long-term impacts. As the authors speculate, these investments may pay off in future consumption in the coming years. The increase in consumption among non-business owners has an even more ambiguous future: if these households used credit to temporarily increase consumption they will have to reduce future consumption to pay down debts. Alternatively, if they used the credit to pay down high-cost moneylender debt, then their current consumption should remain high.

Characteristics with explanatory power are: whether the wife of the household head is literate, whether the wife of the household head works for a wage, the number of “prime-aged” (18-45) women in the household, and the amount of land owned by the household.
6. The economics of saving

A large macroeconomic literature exists to understand national savings rates, their determinants, and implications. Policymakers often strive for target rates, but the variance in national savings rates is remarkable in both developed and developing countries (Gersovitz, 1988; Horioka, 2006). As Rosenzweig (2001) and Collins et al. (2009) argue, focusing on the saving rate places emphasis on asset levels at a given point in time, but that misses the value of savings for many poor households. Many poor households may be actively saving even if their assets at any given moment are low; instead, they are building up lumps of money and spending them within a year. To see this, we need to turn to micro data.

Micro household and individual data reveals much about personal savings rates, the decision-making process at the individual and household levels, and the impact on individuals and households from access to different savings services. We will focus specifically on how informal institutions (e.g., ROSCAs) and formal services influence savings decisions. We will then conclude with a discussion about measuring the impacts from deepening access, both in terms of quantity (e.g., lower transaction costs) as well as the quality of access to savings.

Where households in developed countries, including the poor, may have many products available (savings accounts, automatic transfers, savings bonds, certificates of deposit) to them to help them save and build assets, the poor in developing countries face a much more limited menu of options. Those who are able to save are often forced to invest in risky assets like jewelry or animals or to use informal savings arrangements (e.g., ROSCAs). It is easy to imagine that households in poor countries would save more if they were given access to a broader array of quality savings products. Such access would enable the building of safety nets to smooth shocks and greater accumulation for the purchase of indivisible goods. Much of the discussion on how products influence decisions will build from knowledge and innovation from developed countries, with discussion and examples of applications to developing countries—and a focus on the complexities, risks, and infrastructure unique to the poor in developing countries.

Revisiting “the poor can’t save” assumption

The historical emphasis on credit in the microfinance movement implicitly assumes that the poor cannot save up for investments, and instead must borrow at relatively high interest rates in order to make investments. Criticism of this view is not new, and was argued strongly by the rural finance group at Ohio State University in the 1970s and onward, following McKinnon (1973) and Shaw (1973). The argument has been revisited by some donors, notably the Bill and Melinda Gates Foundation, with a renewed focus on encouraging the promotion of savings in rural settings (Guth, 2008). Adams (1978) decries the lack of focus on expanding rural savings capacity and argues that only a handful of countries stress mobilization of voluntary household savings, arguing that policymakers have too quickly assumed that rural households are too poor to save, that
there is no excess cash flow from income, and that those that do acquire additional income spend it on consumption or ceremonial expenses.

Yet evidence on rural savings behavior from various economies has shown impressive propensities to save among rural households. Adams cites the remarkable savings rates achieved in post-World War II East Asia: the average propensity to save in 1973-74 increased to 0.31 in Taiwan (compared to 0.19 in 1960), 0.22 in Japan (compared to 0.10 in 1950), and 0.33 in Korea (compared to 0.04 in 1965)—a product, Adams argues, of pro-saving public policy. The argument continues that while rural households have a substantial capacity for voluntary saving, their capacity is adversely influenced by rural financial markets that tend to discourage savers through subsidized (but unreliable and unsustainable) credit.

The focus on household decision making is not new either. For example, von Pischke (1978) argues that the rural poor can save, but that they have specific needs due to low account balances, seasonal income, remote location, non-cash assets, and high transaction operations. The development challenge in encouraging savings lies in the design of financial technology to serve these specific needs.

*Problem with concepts of poverty = destitution*

National income accounting and purchasing power parity (PPP) calculations tell us that people in poor countries, on average, do not have a lot of money. But of course what this implies for policy requires much more information. Even for those at the bottom of the income distribution, themselves in countries at the bottom of the global distribution, the fact of earning very little money is insufficient for making reliable inferences about the types of products and services that can be afforded, much less valued.

Take, for instance, Ethiopia, where 23 percent of the population lives below $1/day (PPP) (UNDP). A family of seven at this international poverty line would consume roughly $210 worth of goods and services per month— if they were purchased at US prices. PPP measures are debated (Deaton, 2006), with particular questions about the transferability for the bundle of goods that are purchased by the poor (versus that of the median consumer, for example). However, even if these measures are underestimated by a factor of two, this still implies monthly household consumption of the equivalent of $420 in the United States. These amounts buy so little in terms of meeting an entire family’s needs in a rich country that it can be hard to fathom how they get by at all. Going further requires knowing more about the finances of the poor: what do they spend on food, housing, and healthcare? And do they have any money left over at the end of the month?

Banerjee and Duflo (2007) tackle this question: “how actually does one live on less than one dollar per day,” by assembling a dataset comprised of existing detailed survey data (mostly World Bank LSMS surveys and Rand Family Life Surveys) representing the expenditures of poor households (including households living under $1/day and $2/day) in 13 countries. The surveys were conducted between 1988 and 2005. They find among those living under $1 per day between 56 and 78 percent of household income is spent on
food (slightly less in urban areas). Only around 2 percent of income is spent on education, and a bit less on healthcare. Housing does not seem to be a major expense, perhaps because many very poor households effectively own their own land (though their holdings are quite small).

Though the very poor often fail to meet minimum caloric requirements they choose to make expenditures on many items like alcohol, tobacco, festivals, and radios. As if to reemphasize that these households are truly poor despite their disposable income the authors describe the sample from Udaipur, India, for whom they have the most detailed asset data: among the extremely poor most have a bed or a cot but only 10 percent own a chair and 5 percent a table. Less than 1 percent has an electric fan, a sewing machine, or a bullock cart and no one has a phone. Clearly the very poor are choosing to spend their resources in particular ways. Like households everywhere they have to make choices between consumption, investment, and savings.

Having the right financial products conveniently available might help them make the choices they prefer over the long term. Though some of the surveys predate the global explosion in microfinance, the lack of financial access is nonetheless striking. While borrowing is quite common among the extremely poor (ranging from 11 percent of households in rural East Timor to 93 percent in Pakistan), little of it is conducted with a formal institution. By and large the extremely poor borrow from relatives, shopkeepers, and fellow villagers. The pattern is consistent in recent data from India reported by Banerjee and Duflo (2007), where only 6.4 percent of borrowing is from a bank or cooperative, even when there is a branch nearby. This informal credit is expensive: borrowers pay nearly 4 percent per month. Collins et al. (2009) report similar data in their samples from Bangladesh, India, and South Africa.

Savings is no better. Except for one notable outlier (Cote d’Ivoire, where 79 percent of extremely poor households have a savings account) the fraction is below 14 percent in the other countries (Banerjee and Duflo 2007). That may be even more problematic than the lack of credit since, as the data suggest, even the very poor can use help resisting the temptation to spend money on immediate consumption.

6.1 Basic models of saving

Browning and Lusardi (1996) review nine models used to explain motivations to save: precautionary, life-cycle (to provide for anticipated needs), intertemporal substitution (to enjoy interest), improvement (to enjoy increasing expenditure), independence, enterprise, bequest, avarice, and downpayment. The authors then review the major economic theories of savings, before reviewing empirical evidence on these theories. The data provide a valuable description of who saves and how saving rates have changed over time, but no unique theory explains why people save. We refer the reader to this and other reviews (e.g., Rosenzweig 2001 and Armendáriz and Morduch 2010), and instead here focus on the key motivating and classic models, and then turn to the evidence from developing countries to help understand household decision-making processes.
The lifecycle hypothesis (Modigliani, 1986; Ando and Modigliani, 1963) remains the most influential model of savings. The lifecycle hypothesis (LCH) framework articulates the relationship between consumption, income, wealth, and savings, over the life of individuals. Its central insight is that households have a finite life and a long-term view of their income and consumption needs. They therefore increase their wealth during their working life and use it to smooth consumption during retirement. Wealth itself can come from the accumulation of savings (the difference between “permanent” and “transitory” income) or from bequests. The life cycle hypothesis was one of the first models used to explain savings; it is supported by many empirical analyses in rich countries and is robust to varying assumptions.

As discussed by Deaton (1997), the life cycle hypothesis requires adaptation to fit the multi-generational households commonly found in developing countries. If, for example, the social norm is that a household unit has grandparents, children and grandchildren, the need to save over one’s lifetime is diminished. Rather, intergenerational transfers replace the need for savings and borrowing over one’s life. Naturally, demographic transitions (e.g., AIDS epidemics leading to premature deaths of income-earning generations) can wreak havoc on the ability of intergenerational households to transfer wealth, thus making it important to consider how such trends influence life-cycle savings decisions as the LCH suggests.

For poor households, precautionary savings models are often a better fit (Deaton, 1997). The models capture the fact that for many poor households the volatility of income and the inability to borrow to smooth consumption is potentially just as damaging as a persistently low level of consumption. Rutherford (2000b) puts forward a simple prediction, one that also falls out of most models of savings that generate a preference for smooth consumption: the poor need mechanisms to make small deposits and large withdrawals. The idea is that poor people can save and that they want to save in order to meet life cycle needs, cope with emergencies, acquire assets and develop businesses. Most of these needs come in lump-sums, however, whereas income often comes in little installments (cash labor income, or entrepreneurial income). One exception is agricultural income, with a small number of harvests per year, and we will discuss this below.

Putting these points together: the poor have uneven cash flows (thus the need to save) and they have available income (thus an ability to save). What they require is a safe and convenient place to keep their money and a structure with which to discipline the accumulation of lots of small sums and their transformation into a large sum. Rutherford (2000a) eloquently articulates that the poor can and do save using a variety of homegrown mechanisms including rotating savings and credit associations (ROSCAs) and deposit collectors who charge the poor to take their savings. Each option presents tradeoffs in convenience, risk, price, and simplicity. The importance of the last dimension, simplicity, should not be underestimated: as programs and products get more complex, they have a greater potential to meet the specific financial needs of the poor—but they also get become harder for clients to understand and manage. Good research can help programs identify optimal combinations of all these dimensions by determining
which factors “sell” the best to clients and by providing an understanding of the literacy necessary for a product to succeed (both in terms of take-up and proper usage). However, as we discuss below, even the “best” combination of convenience, security, and price may not be enough: just because people make decisions does not necessarily mean they make good decisions—and by “good” we mean decisions about actions that people say they want to make — like building a safety net for hard times or saving up for profitable investments.

Features such as transaction costs, liquidity, and interest rates influence the take-up and usage of financial services. But other factors matter, perhaps in some cases even more. In an experiment in South Africa, Bertrand et al. (2010) tested the relative importance of interest rates, marketing features, and choices on a direct mail solicitation to consumer borrowers from a regulated for-profit microlender. They found that interest rates mattered, but that simple marketing ideas mattered even more. For instance, giving consumers only one choice on loan size, rather than four, increased the take-up of loans just as much as if the lender reduced the interest rate by about 20 percent. Even more striking, they found that if there was a photograph of a woman on the direct mail solicitation, this drove take-up, among both men and women, of the loans just as much as if the lender reduced the interest rate by about 33 percent. These are large effects, particularly given the industry focus on “important” characteristics, such as price.

What does this tell us? We need to pay attention to more than just the pure economics of the choices being offered. The way offers are presented can have just as much to do with take-up and usage as do the terms of the account. Recent work by Richard Thaler and Cass Sunstein has documented a plethora of examples of “choice architecture” in which the way choices are presented and structured may matter just as much, if not more, as the choices themselves (see Thaler and Sunstein (2006), and a popular press book, Nudge, (2008)).

41

Given that the presentation of choice matters, targeted research can help answer the question of how products and processes can be designed to most effectively assist poor people to accomplish their stated goals. Are there certain temptations, for instance, that individuals would actually prefer not to have? For example, do individuals find they purchase items, such as entertainment goods, that they later regret having purchased (when, for example, a health shock occurs)? Commitment savings accounts can help prevent consumption of goods that are later regretted. Altering “default” settings is another example of how product design can nudge individuals towards decisions they say they prefer. In most decisions, something must be established as the action that occurs if no alternative action is taken. For example, should a proportion of a remittance or paycheck be set aside automatically into a savings account?

In order to understand how product design might matter, we need to first understand something deeper about the psychology of the decision-making process and the household and societal constraints that lead to certain decisions and coping strategies. Furthermore, employing one design over another requires recognizing that there is rarely a “neutral” option: choices in product design will affect how the product is used, and by
whom. Some users may benefit more than other in particular ways. As such, every choice about how to design and offer a product requires taking a normative stand on what will improve outcomes for individuals.

6.2 Constraints to saving

If people do not save as much as they wish they would, why not: what are the constraints? We think about these as demand-side and supply-side constraints, although clearly there is a relationship between the two.

It is useful to first note the high price the poor are willing to pay for savings services that they can trust. This shows clearly the high value they place on savings more generally, and thus the welfare improvement possible from identifying policies that provide savings options for the poor at lower costs. For example, a ROSCA is a common method of savings that provides zero interest income, and significant loss risk. Wright and Mutesasira (2001) document these risks for informal savings. Through a study of MicroSave data, they indicate that people with access to the formal sector reported much higher savings than those without access. The percent of those reporting losses from formal mechanisms (15%) was much lower than reported losses in the semi-formal (26%) and informal (99%, albeit loosely defined) sectors. On average, clients in the informal sector lost 22% of their savings. Note that this includes informal savings such as livestock, which can die but also provide a potential return if they live. Thus this is not the ideal analysis to show that, holding returns constant, informal savings are riskier. In this light, the issue of allowing microfinance institutions to offer savings products should be considered in terms of relative risk, since savings are relatively safer at microfinance institutions than in informal mechanisms. Rather than dictate the decision of where the poor save, it is key to help them make informed choices by helping them understand the relative risk of semi formal institutions.

The popularity of deposit collectors also offers clear evidence of the demand for savings, and offers insight into some of the potential behavioral constraints on savings. In Ghana, for instance, individuals pay for informal deposit collection services through susu collectors who travel to individual homes or businesses at regular intervals in order to collect savings deposits. In some cases, payment for this service is high enough that individuals’ real return on savings is negative. Aryeetey and Steel (1995) document this and discuss the basic structure of these services. The susu collectors collect deposits from customers (mainly women operating market stalls) every day and return the accumulated savings at the end of the month. They neither pay nor charge interest on the deposits but they keep one deposit per month as a fee. This 3.3 percent fee works out to a negative annual return of 54 percent for those who withdraw their deposits at the end of each month.10

10 Despite the high cost Susu services are extremely popular: Aryeetey and Gockel (1991) reports 78 percent of market women in Ghana’s largest cities using Susu collectors to save. Susu collectors lend too, but the risk to depositors seems limited: the collectors surveyed by Aryeetey and Steel lent only small amounts to a small portion of their client base. Borrowers are screened based on regular savings history,
These issues lead to a re-evaluation of the two views that dominate thinking around the financial behavior of the poor (Mullainathan and Shafir (2008). One view positions the poor as rational individuals who are methodical and calculating in their financial decisions; the other positions them as impulsive and misguided. Mullainathan and Shafir present an alternative and more realistic perspective, that the poor are neither completely rational nor irrational. Just like everybody else they make good decisions some of the time, and rash or uninformed decisions at other times. They are subject to the same psychological biases as the wealthy; the main difference is that these behaviors have a more profound effect on the poor because of their narrower margins of error due to their adverse financial condition. For example, if a poor individual makes a mistake, it may lead to their telephone being cut off (or in a developing country, to not having sufficient funds for buying “load” on their cellphone). Not having telephone access may then lead to a problem at work, in which the worker is unavailable to call in. This then leads to loss of income, which then leads to further problems. A wealthier individual may have made the same initial cognitive error, but it simply did not reverberate and escalate to further problems in that person’s life.

Putting psychology into conversation with economics changes the way we think about some basic relationships. Despite a preponderance of arguments for lowered transaction costs as a panacea to savings policy (Robinson, 2001), for example, behavioral theories suggest that higher transaction costs can, in some cases, be welfare enhancing. The problem lies with limited commitment problems, in which individuals want to make certain future decisions but face constraints from their own weaknesses, or constraints from others. Beginning with Strotz (1955) and Phelps and Pollak (1968), theoretical models have been put forth that describe outcomes that arise when there are inconsistencies between current temporal tradeoffs and future temporal tradeoffs. These models often incorporate hyperbolic or quasi-hyperbolic preferences (Ainslie, 1992; Frederick, et al., 2001; Laibson, 1997; O'Donoghue and Rabin, 1999), theories of temptation (Gul and Pesendorfer, 2001; Gul and Pesendorfer, 2004), or dual-self models of self control (Fudenberg and Levine, 2005) to generate this prediction. They all share one consistent prediction: individuals should exhibit a preference for restricting their future choice set, and being able to do so will increase their ex-ante welfare. Naturally, this requires that individuals are self-aware enough to volunteer for such restrictions (much like Odysseus tying himself to the mast to avoid the tempting song of the sirens).

Laibson (1997) looks at decisions made by a consumer with access to illiquid assets – these are assets that generate substantial benefits in the long run, but no immediate benefits (“golden eggs”). The theory suggest that financial innovation, which increases liquidity and reduces implicit commitment opportunities, could have led to the lower savings rates in the US. The model implies that financial market innovation may reduce

and the Susu collectors typically lend only half of a month’s deposits to a given client. On average less than 10 percent of their portfolios were exposed to lenders.
welfare by providing “too much” liquidity. We revisit the implications of this work below when we discuss technological innovation that lower transaction costs for savings.

Evidence of such preferences are often difficult to show, but a growing literature is demonstrating that demand exists for such restrictions, and in some cases firms are responding by offering such services. In the United States, Christmas Clubs, popular in the early 20th century, committed individuals to a schedule of deposits and limited withdrawals, typically with zero interest. In more recent years, defined contribution plans, housing mortgages, and tax over-withholding now play this role for many people in developed economies (Laibson, 1997). For example, on the Earned Income Tax Credit, many individuals do not take advances, effectively an interest-free loan to the government, potentially as a costly commitment device to save (see Jones, 2009).

In developing countries, informal institutions have arguably played this role for years. Many theories exist to explain the presence and structure of rotating savings and credit organizations (ROSCAs), but one commonly held belief is that they provide a form of commitment from your future self (Gugerty, 2007) or from your spouse or extended family (Anderson and Baland, 2002). In Gugerty’s work from Kenya (2007), qualitative and quantitative evidence from 70 ROSCAs in rural Kenya was consistent with a self-control commitment story: nearly 60 percent of the ROSCAs had an explicit spending agreement, for which members were required to identify in advance their purpose for the pot of funds when their turn is up. Members would verify each other’s purchases to ensure they adhered to the intended purpose (though most used the money for more than one purchase). Drawing causal conclusions from the data is difficult because in practice new members had limited choice in selecting a ROSCA. While spending-agreement ROSCAs show higher savings rates, they also tend to be composed of wealthier members, often with formal-sector income. Much of the qualitative evidence revealed individuals discussing inability to ‘save alone.’ Married women were not more likely to participate than non-married, nor were women whose husbands lived on their compound more likely to join. This suggests that the ‘save alone’ reference is fundamentally about one’s own ability to save, or about claims made by non-spouse family or neighbors.

ROSCAs also have been shown to be spousal control tools, not just self-control tools. In Anderson and Baland (2002), we learn from Kenya (although a more urban setting than the Gugerty paper discussed above) that women with some but not all of the household income are more likely to participate in a ROSCA than those with all or none. The paper works through the following model. It assumes a husband and wife differ only in their preference for an indivisible good, for which the wife values the good, and the husband does not. Assuming that the husband will respect the ROSCA institution (i.e., a man cannot punish a woman for joining the institution, presumably due to a societal norm), then women who need to extract funds from the household in order to save up for an indivisible good will find the need to join ROSCAs. Women who earn all of the income presumably have power, and thus under this model, no need to join a ROSCA.11 Women

---

11 Two alternative stories are worth mentioning. Women who have all of the income may also be too busy for the ROSCA, if the ROSCA is time consuming, and women who earn all of the income in the household
who have none of the household income, on the other hand, have no control over any income flows, and likewise are unable to join ROSCAs because they are unable to commit to any future deposits. This parabolic relationship is exactly what Anderson and Baland find when trying to predict which type of wives join ROSCAs. Naturally, these two stories are not mutually exclusive, and commitment devices in general can work for very different reasons. Understanding the relevant importance of these models may have important policy implications, such as how to design marketing and account access rules.

The need for commitment from others is not limited to spouses. Platteau (2004) examines the way in which egalitarian norms may inhibit personal savings, and thus growth. In some societies successful individuals are called upon to transfer some of their own wealth (directly and indirectly through transfers and favors) to benefit poorer community members and kin. In cases where the wealthy do not consider these to be legitimate contributions, this even serves as a disincentive to work hard. These individuals may also resort to a number of alternate strategies to hold their wealth that involve high transaction costs, in an effort to reduce demands on their income and wealth. These behaviors result in inefficiency that inhibits economic progress and reduces saving levels.

Whereas these studies argue that indigenous institutions have been formed in order to satisfy a demand for commitment, they leave open the question of whether stronger, more formalized commitment devices could succeed in both attracting the right individuals and helping them to adhere to their stated preferences. In the context of Thaler and Sunstein (2008), formal institutions can “nudge” individuals via product framing and design towards decisions they claim to want. For example, Ashraf, Karlan and Yin (2006b) designed a commitment-savings product, called SEED, that provided clients with a commitment to not withdraw their funds until a goal was reached. SEED clients voluntarily restricted their right to withdraw any funds in their own accounts until they reached a self-specified goal. Clients could opt to restrict withdrawals until a specified date (e.g., in a month when school fees were due), or until a specified savings amount was reached (e.g., a certain amount of money for a new roof). The clients had complete flexibility to choose which of these restrictions they would like on their account. However, once the decision was made it could not be changed, and SEED clients could not withdraw funds from the account until they met their chosen goal amount or date.

To evaluate the impact of this new product the bank implemented a randomized control trial where it assigned individuals to either receive an offer to open the SEED account or not, or a third group which received a marketing pitch about the importance of setting goals for savings (but no access to a new commitment savings account). Among those offered the account, 28 percent opened one, and, importantly, those who exhibit inconsistent time preferences in survey questions about hypothetical alternatives were more likely to open an account. After twelve months, average balances increased by 80% may also be better financial managers, less likely to have self-control problems, and thus also less likely to join a ROSCA.

Interestingly, this preference for commitment was strongest amongst women. The study did not provide data to help understand heterogeneity across gender, and thus future work on this issue would be fruitful. It
percent in the treatment group (i.e., those who got the SEED offer) compared to the control group. The account offer was also associated with a significant increase in women’s decision-making power within the household (as measured both qualitatively by asking who has power over certain decisions, as well as more objectively by observing the gender-bias of durable goods purchased) (Ashraf, et al., 2010).

A question remains, however, to what extent such commitments are about binding one’s behavior, or are in fact merely about creating structure. The previous section described how those wanting to save might opt for credit simply because it provides a defined schedule for making regular deposits. Karlan et al.(2009a) test the effects of simply making savings more salient by sending clients simple reminders to make deposits. They find even with no commitment, the reminders can be successful in increasing savings rates (by 6%) and helping clients meet savings goals (a 3% increase in the likelihood of reaching one’s goal). Similar positive impacts on savings were found by a deposit collection services tested in Ashraf et al (2006a), as well as Dupas and Robinson (2009).

These ideas are gaining acceptance within academia. They are also having impacts in practice. Grameen Bank of Bangladesh, for example, has launched a successful commitment savings device, a “pension” product that requires monthly deposits in fixed amounts and returns savings (with interest) after five or ten years (depending on the product). The savings account is a “pension” in name only, and while it is used to pay for old age, households also use it to accumulate for housing improvements, wedding expenses, migration and the like (Collins et al. 2009, chapter 6).

Planning and financial literacy

A preponderance of evidence shows that financial illiteracy is prevalent around the world, and is correlated with low savings. However, this begs the policy question of whether interventions intended to increase financial literacy can in fact lead to changes in behavior of importance. Evidence is clear that people, everywhere, are financially “illiterate” by many definitions, both described and measured by lack of basic numeracy (e.g., simple compounding), financial knowledge (familiarity with financial products, including credit, savings, and mortgages), and financial planning (e.g., saving for retirement). Lusardi (2007) shows that financial illiteracy even in the US population is widespread, and particularly acute for specific demographic groups, such as those with low education, women, African-Americans, and Hispanics. Moreover, close to half of older workers do not know which type of pensions they have and the large majority of workers know little about the rules governing Social Security benefits. Notwithstanding the low levels of literacy that many individuals display, very few rely on the help of experts or financial advisors to make saving and investment decisions.

Naturally, the mere correlation of financial illiteracy with outcomes such as savings decisions does not imply that financial illiteracy is the cause of the low savings. Several endogeneity issues could be at play, including the simplest, that omitted variables such as

is suggestive of an important interaction between gender and the preference and need for commitment contracts, as discussed here.
motivation to succeed are the true cause of both financial illiteracy and low savings. Or, reverse causality: if one is incapable of saving in quantity then one would be unlikely to invest in knowledge of savings vehicles. Examining the causal impact of financial education programs requires effective methods for establishing the counterfactual.

Two studies on financial literacy examine this very issue, and draw strikingly different conclusions. This demonstrates the difficulty of establishing attribution in evaluation of public policies. The first paper, Bernheim, Garrett and Maki (2001), uses variation in state-mandated financial education to measure the treatment effect of financial literacy training on household savings. Using data from Merrill Lynch, and a telephone survey of 3,500, the authors employ a difference-in-difference approach and assume that timing of the introduction of state-mandated financial education is exogenous. They conclude that the mandates were effective in teaching basic financial literacy, and led to a 1.5 percentage point higher saving rate. However, it turns out there is evidence against their key identification assumption. Cole and Shastry (2008) extend this study with more data and conclude that the Bernheim et al result was spurious. Census data, and thus a larger dataset, allows for the inclusion of state fixed effects to control for unobserved, time-invariant heterogeneity in savings behavior across states, as well as non-parametric identification of the treatment effect itself (rather than a linear measure of years-since-mandate-began employed by Bernheim, Garrett and Maki). Once these three enhancements are implemented, all treatment effects fall to a precisely estimated zero, thus both eliminating the conclusion that financial literacy as implemented under this program had any effect, and demonstrating a key endogeneity issue that plagues this literature (that rollout of programs is responsive to demand and thus extensive work must be done to create convincing counterfactuals).

Savings, as we have seen, is important to poor households, important enough that they are willing to pay for the service. This commitment to save has led some practitioners to conclude that savings, rather than credit, is the more practical strategy to promote, especially for the very poor. This is surely true for some households, but is strongly contested by the evidence in Collins et al. (2009).

There are some settings where the poor do not save, even where it is clearly advantageous for them to do so. Why haven’t more households saved their way out of credit constraints? Such puzzles suggest that large-scale savings promotion may require a more sophisticated strategy than imagined initially: it will involve better research to understand why the poor do not save when they could, and which strategies can help overcome these barriers.

A related puzzle concerns entrepreneurs who borrow persistently, not for a one-time business expansion but for routine working capital. This is very expensive. In extreme examples vegetable vendors in India are known to borrow small sums each day to purchase vegetables, repaying each afternoon from their daily sales. They pay rates as high as 10 percent per day. Among a sample of vendors in Chennai, 50 percent claimed to have been engaging in this type of borrowing for at least ten years! What is so interesting about such borrowing cycles is that it is easy to show that in principle by
saving merely one rupee (a few cents) each day (and borrowing that much less) the vendor could be debt free, and able to finance her own working capital, in just 50 days. From there her returns are enormous: her daily profit margin is boosted by the ten percent she was paying to the moneylender. So why do the vendors persist in borrowing at such rates? The answer remains unclear but financial education seems a good place to start. If they understood the true cost of their borrowing, they might well be convinced to save.

In India and the Philippines, Karlan and Mullainathan (2008) explore the impact of debt payoff and financial education on the persistence of high interest debt. These experiments address the aforementioned endogeneity concerns by randomly offering indebted vendors different experimental treatments. Specifically, some vendors are offered an unexpected debt payoff, others financial literacy training emphasizing the utility of savings to finance business expansion and the cost of debt (the training itself was a brief but focused 30 minute session), and a third group are offered both debt payoff and financial training. A fourth group serves as a control group and thus received no experimental treatments. The evidence from the Philippines suggests that debt payoff can have immediate but then dissipating effects on vendors’ reliance on high-interest loans: vendors who received one of the payoff treatments were 31 percentage points less likely to have money lender debt almost one month after the payoff and 17 percentage points less likely to have taken a high-interest (defined as greater than 5 percent per month) loan after three months. Preliminary results showed that the effect dissipates over time, however, as individuals gradually go back into debt.

Of course a measured reduction in household consumption is not necessarily a good thing. More research will be required to look deeper into the specifics of how household consumption decisions change as a result of greater financial literacy. Are households choosing to save rather than spend excess income? Are they making immediate sacrifices in the interest of financing longer-term business goals? Another question that remains unanswered is whether the impacts observed in the debt payoff and financial literacy training experiment can be multiplied by providing high-interest borrowers with an effective savings mechanism. These and other questions require further research to disentangle how product offerings can change the cash management and thus savings and consumption decisions of the poor.

**Pricing**

Just as we discussed earlier with respect to credit, many have assumed that because the poor are willing to actually pay to save, this implies that their demand is inelastic with regard to price. Naturally one does not imply the other, since willingness to save at zero or negative real interest rates simply means that at that price demand is positive, but at higher prices of course demand could still be higher. Hirschland and Owens (2005) provide a useful overview of a typical practitioner's perspective on the considerations for how to set price for savings, including competitive analysis and cost drivers. Clear information on elasticities, however, would help tremendously.
In Karlan, Mullainathan and Zinman (2008b), just as in earlier work on credit elasticities, the authors conducted a series of field experiments in which the interest rate on the savings account was randomized, and data collected on how this influenced the decision to open an account as well as the volume of savings held in the account. Specifically, the treatments tested are high, low, and “reward”, in which the high rate was 1.0 percent per annum higher than the low, and the bonus was also 1.0 percent but only awarded if the individuals savings goal amount was reached by their goal date. In the second experiment, the design was slightly modified, and the increase in interest rates and bonus was 1.5 percent per annum instead of 1.0 percent. While they find that the product take-up rates among the high-interest group are higher than that of the low-interest group in both experiments, and both lead to 3 percent higher savings balances, still after 7,000 offers (with an average take-up rate of 23 percent and average balance of about US$10), neither of those results are statistically significant.

Loss Aversion and Mental Accounting

The concept of loss aversion is well established in our understanding of human behavior but not well adopted in the design of savings products. The 2002 Nobel Prize in Economics was awarded to Daniel Kahneman for, among other ideas, the simple demonstration of this. The canonical classroom experiment involves mugs (Kahneman, et al., 1990). Half of the students are randomly selected in a lottery to each get a mug. Everyone in the class is then given the opportunity to trade. In an “efficient” world of no loss aversion, the half with the highest valuations should end up with the mugs, irrespective of whether or not they “won” the lottery. But winning the lottery changes one’s valuation of the mug. Why? Because once the mug is won, the reference point is shifted. Giving up the mug is now a loss, whereas, for everyone else acquiring the mug is a gain. Losses loom larger than gains, and few people trade. Those who won the mugs mostly fail to find someone willing to pay them enough to make them part with their moments-ago-won mug. Instead, if the professor asked for students to announce their valuation before receiving the mugs, then handed them out and made everyone adhere to their stated valuations, one would expect, on average, half of the mugs to be traded.

Now let’s apply loss aversion theory to the decision to make a deposit into a savings account. What is a savings deposit in a mental accounting system? It is trading off some current (salient) consumption for some future (nebulous) consumption. Ignore the timing for a moment (more on that below), and what does one have? A sure loss, in exchange for an unclear gain. Thus loss aversion may drive the individual to consume, rather than save.

How does one tackle this? Can product design effectively convert the savings deposit into a gain? Some of the problem is driven by the vagueness of the future gain. Perhaps making the future gain more salient, one can convert the savings deposit into, at a minimum, a “neutral” and, at best, a “gain.” Ashraf, Karlan and Yin (2006), in the study in the Philippines discussed earlier, employed a “placebo” treatment involving marketing only, in which bank marketing agents visited homes of prior and current clients to encourage them to consider setting savings goals. This treatment was intended in that
study as a placebo, in order to make sure the treatment effect from the commitment savings product was due to the mechanism design of the product, and not the labeling and goal setting promotion. In fact, the impact of the goal-setting treatment (referred to there as “marketing” treatment) was positive, although not statistically significant. Labeling can also matter. In developed countries, we have seen accounts “labeled” through marketing (and tax advantages): savings accounts for education, savings accounts for health, savings accounts for retirement. These ideas do not require fancy infrastructure, but rather mere marketing and packaging. Does account labeling allow one to make the later gain from saving more salient in the present, and thus cancel out the loss from foregoing current consumption? This is an open question.

More generally stated, one could imagine many methods for converting the “loss” of the foregone consumption due to making a deposit be converted to a gain. Doing so relies heavily on the lack of fungibility of money, a violation of most traditional models in economics. Mental accounting, as put forward by Thaler (1985) provides a framework to interpret such behavior.

Default options

Extensive evidence exists to show that default option matter. This is true in many facets of life, and savings behavior is no exception. Madrian and Shea (2002) document this clearly for individuals making retirement decisions in the USA. There, Madrian and Shea show that setting automatic enrollment as the default participation in retirement plans leads to a 50 percentage-point increase in the likelihood of participation due to automatic enrollment, and similarly large effects are found on the type of investments individuals choose.

A more complex example comes from Thaler and Benartzi (2004). Here, the authors implemented a program called “Save More Tomorrow™” (SMarT) in which individuals weakly commit (by “weakly”, we mean they can reverse this decision at any time) a portion of their future salary increases toward retirement savings. The paper cited above reports findings based on evidence from the first implementation of the SMarT program, through four raises. Key findings are (1) a high proportion (78 percent) of those offered the plan joined, (2) the majority (80 percent) remained in the program through the fourth raise and (3) the average savings rate increased from 3.5 to 13.6 percent over 40 months. This program has now been adopted for many corporate retirement plans in the United States. In a survey of 146 employers in the United States, Hewitt Associates found that 31 percent of plans have an automatic escalation program, and of the 69 percent that do not, 42 percent said they plan to incorporate such a component into their plan (Hewitt, 2007).

Adapting these ideas to developing countries has tremendous promise, and also could provide ample opportunities for learning why this works more precisely, as several factors were employed at once in the design of SMarT. First, since upon sign-up, the default future decision would be to save more, action had to be taken to change this decision. Second, by framing the future increase as “coming out of your next pay
increase” (even if the next pay raise was merely an inflation adjustment), money illusion may have led more to sign-up than would otherwise have occurred. Third, the method of presentation, and the skills of the advisors, may have influenced the decision of individuals to participate.

Marketing

Often design and discussion about products by academics focus on the terms and structure and risks of a product. Yet, in many cases, the presentation, framing, and promotion style influences the outcomes. For example, in Landry et al (2006), in a door-to-door marketing of a fundraising appeal, the authors found that the physical attractiveness of the door-to-door salespeople was far more important than the lottery that was being offered to some but not others. Similarly, in developing country setting, Bertrand et al. (2010) find that adding a photo of a woman to a direct mail solicitation increases the likelihood of borrowing by just as much as dropping the interest rate by about 30%. Al-Bagadi and Cracknell (2005) discuss this more generally in the context of marketing and promotion of savings in developing countries. They argue that microfinance institutions must translate ideas about why potential clients should want to deposit with a particular institutions into a message that motivates them to do so. We suggest that future studies that examine product features as discussed above do not forget to think about the promotional strategy, and ideally incorporate the promotional strategy as an integral part of any study.

Social networks and peer effects

If the poor have limited knowledge of the benefits of saving, then learning about these benefits through social networks could potentially have important impacts on the poor’s savings behavior. Despite its relevance, not much research exists on the role of peer effects in savings decisions. Peer effects are notoriously hard to identify empirically. In many instances individuals’ decisions within a social group are correlated for reasons that have nothing to do with social learning or social imitation. Behavior of individuals may be correlated simply because individuals in the same group have similar unobserved characteristics – e.g. a common propensity to save – or share a common environment. Manski (2003; 2005) lays out some of the difficulties involved in identifying peer effects.

In an empirical contribution, Duflo and Saez (2002) investigate the role peer effects play in retirement savings decisions in the United States. They study whether the decisions of employees of a large university to enroll in a university sponsored Tax Deferred Account plan are affected by the decisions of colleagues in the same department. Instrumenting average participation within peer groups by average salary or tenure and looking at subgroups of peers within departments, Duflo and Saez find evidence that the individual participation decision is influenced by the decision of one’s peers.

Assessing the importance of peer effects in savings decisions matters for the design of policy interventions. If innovations in savings behavior spread through social networks, the impacts of financial education efforts can be much larger than the impacts on the
financially educated. An initiative that aims to make effective use of learning from peers to promote savings is Oxfam’s Saving for Change program, based on a model used in Nepal and now replicated in Africa. Saving for Change groups are informal self-managed saving and credit groups consisting of 25-30 women. Group members meet weekly to save a predetermined amount and the collected funds are used to make loans to group members. Since loans are repaid with interest, the fund grows over time. At the end of the cycle the fund is divided according to each member’s share in the savings. The program has been found to rapidly reach considerable numbers of women. The first group in a village is typically trained by an external agent with the members of the first group then forming and training subsequent groups in the village. The role of social networks and peer effects in the spread of these saving and loan groups is the object of ongoing research.

Peer influences need not work merely through one-on-one interaction. For example, as Garon (2004) points out, East Asian states played a key role in inculcating savings habits and thrift amongst their citizens. Asian values of savings and consumption were created by a regionally adopted model of state promotion of savings, with Japan's colonial presence in these countries acting as one of the major catalysts to these state efforts. The history of the "Japanese Model" includes national campaigns to promote savings, postal savings banks, and a Central Council for Savings Promotion. This model was exported to varying degrees to South Korea, Singapore, and Malaysia. Garon argues that though thrift and savings are not timeless or unique Asian values, many Asian people have come to embrace these as part of their national identities. This begs the question, though: are savings levels cultural, driven by social norms and macro-level policies, or simply the product of one huge omitted variable, e.g., institutions.

Technology

There is a massive effort underway to harness technology to overcome the transaction costs that have prevented the poor from accessing banking services, especially in rural areas. Technology also appears to be revolutionizing the remittance industry, as technology solutions are making remittances easier and cheaper than ever before. As discussed earlier, remittances historically, and still today in many countries, are expensive as a proportion of money sent, particularly for small amounts. Yet unlike credit there are no information asymmetries or default risks to explain these costs, and hence technology solutions are proving to have big impacts on the costs consumers are offered.

The donor consortium CGAP has launched a campaign to bring mobile banking services to 25 million low-income people by 2012. It is early to tell what form these services will ultimately take, but right now attention is focused on point-of-sale devices installed at retail agents (e.g., local stores) who can accept deposits or payments, as well as withdrawals; and banking via mobile handsets, enabling customers to transfer funds electronically. The explosion of accessible technology, no doubt, will change whether and how people interact with banks. This may have important unintended consequences. As we discussed above, liquidity can actually be undesirable for some. Could electronic
banking lead to further problems with self-control and spousal, familial, or community pressure? If hidden savings are not offered in conjunction with increased access to funds through technology, such pressures could lead to worse outcomes. As transaction costs drop, this will become an even more important area for future research and innovation.

Furthermore, much of the advent of microfinance has been through innovations that remove layers of information asymmetries so that banks can profitably lend to the poor (and for microinsurance, so that they can insure the poor). Removing the human touch from banking may have unintended adverse consequences, in that individuals will reach for the “easy” technology but in the process lose the interaction with the credit officer that is necessary for the financial services firms to establish relationships and lend and insure. The personal relationship with a banker may reduce information asymmetries both by allowing the bank to have more information for screening, and also by reducing moral hazard, by instilling a sentiment of reciprocity or personal loyalty in the mind of the borrower, making them less likely to engage in moral hazard.

Thus, technological innovations in the user-interface are promising for radical reductions in cost of accessing savings services, but we need to understand better how behavior will change as a result of this easier access. In a quasi-experiment from the field, albeit not in this context, consumption on a particular good (sorting laundry into two washing machines or pooling into one) increased by 50% as the payment mechanism shifted from cash to prepaid cards (Soman, 2003). How will mobile “minutes” be treated, as cash to save or money to spend? This simple mental categorization could have serious implications for the long-term impact on consumption and savings from such technologies.

Early experiences with branchless banking are already showing clients are using the new technologies largely for payment transactions, rather than savings or credit. To some extent this may be the result of marketing campaigns by mobile operators which focus on the transfer of money rather than the storage of money—apparently because the operators are concerned about appearing to market themselves as banks and thus attracting notice by regulators. Despite the great potential for branchless banking to expand financial access, thus far mobile banking customers in developing countries have been wealthier customers in urban areas (Ivatury and Mas, 2008).

**6.3 Impacts of saving**

The impact of savings programs can be difficult to measure, both because savings is hard to capture in survey data and because it is hard to isolate savings from other financial services: few institutions offer only savings. Three factors particularly complicate measurement: size, timing, and diffusion. Unlike credit inflows, which can be sizable relative to household income, savings flows can be quite small, and balances accumulate slowly. Also, the timing of the change in behavior and outcome is less clear. For households, savings develop slowly through a small reduction in consumption over time, with a large inflow later. At some point the household will have built up enough savings
to protect themselves from shocks (like sickness or unemployment), to pay school fees, or to start a business. But when is that point? It may not be a simple question of waiting for savings to accrue: household cash flows may vary over time. Researchers need to measure savings balances at multiple points in time, often over several years.

The last issue, diffusion, is perhaps the most important. Poor households save in all sorts of different ways. In addition to a formal savings institution they may save cash at home, with deposit collectors, within a savings club (such as a ROSCA), or by lending to family members. They may also save in non-cash assets such as jewelry or livestock. An evaluation that failed to capture these different savings vehicles could understate savings rates, potentially leading to incorrect conclusions about the impact of the program. Or, an evaluation which focused on just one savings channel may incorrectly conclude that net savings goes up, whereas in fact the only impact was a shift from one type of savings to another.

As with credit, it is problematic to compare savers to non-savers. Savers may be better educated or have more disposable income. They may be better planners, or more risk averse. There might also be reverse causality: healthy people might be able to earn more, or spend less on hospital visits, increasing savings balances. Given these issues it is unsurprising that there are few rigorous evaluations of savings. Burgess and Pande (2005), described in the previous section, solve the identification problem, showing that financial access reduces poverty, but they are unable to separate the effect of savings from credit.

Aportela (1999) evaluates the impact of a government savings program in Mexico, Patronato del Ahorro Nacional (Pahnal). In 1993 Pahnal decided to expand its savings operations through the postal office network. By using the postal offices, Pahnal was able to locate close to many people who were otherwise inaccessible, but without the fixed cost of opening a vast network of offices. As part of this expansion, Pahnal offered two savings options: (1) a fixed-term instrument (12, 24 or 36 months) with compulsory monthly deposits of five dollars, and inability to withdraw until maturity, and (2) a liquid savings account, with minimum balance of just over five dollars, with no fees but lower interest than the fixed term. Aportela uses the partial expansion of the program to compare the change in outcomes for those in communities that received the Pahnal expansion to the change in outcomes for those in communities that did not receive the Pahnal expansion. The analysis uses data from the 1992 and 1994 Mexican Household’s Income and Expenditure surveys. The advantage of these data, rather than using client data, is that there is no individual selection bias. Given that the program chose its own expansion path, however, it is critical to verify that preexisting savings rates or poverty rates are not correlated with the treatment communities. Aportela finds no correlation, nor any evidence of an operational plan that could have confounded the analysis (e.g., an initial expansion into districts that had expressed high demand for the program). The expansion appears to have more closely followed Pahnal’s operational convenience i.e., proximity to its prior branches. Hence an issue remains as to whether Pahnal’s prior branches were located strategically such as to create a preexisting trend and thus selection bias.
The impact on formal-sector savings balances was noticeable: the average savings rate increased by 3 to 5 percentage points. For low-income individuals it was even higher: an increase of 5.7 to 8 percentage points. The emphasis of the analysis in the study is on formal sector savings volume, broken down by different income levels. The attempt to measure aggregate savings, including informal savings, fails to reach decisive conclusions because of lack of statistical precision and data limitations: the analysis is unable to rule out substitution from other savings vehicles. Moreover, it does not allow one to isolate which of the two products, or within either product which particular feature, led to the impact. Whether the increase in formal savings represents an increase in net savings or a shift from informal to formal savings it can be considered a positive impact in that the formalization of savings implies safer savings for individuals.

Impact from specific product trials is more limited. The evidence discussed earlier, from Ashraf, Karlan and Yin (2010), showed a potentially good relationship between savings and empowerment. The commitment savings product there led to a significant increase in women’s decision-making power within the household (measured by an index of decision-making authority over various types of purchases, as well as family planning and children’s education), and an increase of the purchase of female-oriented durable goods. The impacts were particularly strong for women who have below median decision-making power in the baseline data.

Dupas and Robinson (2009) worked with a community bank in rural Kenya to provide incentives to open a savings account to randomly selected entrepreneurs, for whom the researchers paid the fee to open the account and provided the minimum account balance. The control group received no incentives but were not barred from opening an account. The incentives were strong enough that 89 percent of the treatment group opened an account while only three individuals in control group did so. The researchers find remarkable impacts despite substantial transaction fees charged by the bank ($0.50 or more) and the fact that many never used the account after opening it. In contrast to the Karlan and Zinman (2009a) study of the impact of credit in the Philippines and de Mel et al. (2008b) study of returns to capital, here the impacts are found only among female entrepreneurs. Four months after opening the account women show 40 percent increases in productive investment, and after six months daily consumption is approximately 40 percent higher than the control group. In this study however women have different enterprises than men, and hence the gender difference may be properly interpreted as a difference generated by occupational choice. Further research and expansion can help understand this important result.

7. Risk Management and Insurance

To be poor in most of the developing world is also to be disproportionately vulnerable to risk (Dercon, 2004a; Morduch, 1994). Among sources of vulnerability are the high correlation of poverty and ill health (e.g., Case, et al., 2002; Dercon and Hoddinott, 2004), the riskiness of agricultural occupations (Dercon, 2004b), employment instability
within the informal sector (Lund and Nicholson, 2004), and the broad insecurities that arise from weak legal protections (e.g., Field, 2007).

One accounting of links between poverty and vulnerability is offered by Dercon (2004b, Table 1.1), who reports on a survey of rural households in Ethiopia who were asked to look backward on the two decade span 1974-1994. Most households reported having suffered seriously due to harvest failure brought on by drought, flooding, frost, or pests (78 percent). Others seriously suffered due to illnesses and death, both of family members and of livestock (about 40 percent). Like the harvest failures, the health losses are potentially insurable. Other serious losses suffered, however, arose from forces that would be near impossible to insure with traditional insurance products, notably policy shocks resulting in forced labor, new taxes, and migration bans.

Collins et al (2009, Table 3.1) similarly report on major losses suffered by small samples of poor and “near-poor” households during a single year in rural and urban Bangladesh, India, and South Africa. Half of the Bangladesh sample suffered due to serious injury and illness during the study year, as did 42 percent of the India sample. These health crises were accompanied by losses to income and property. For example, in Bangladesh, the urban individuals suffered from slum clearing by police and construction workers, and in rural India individuals suffered from a particularly bad harvest. When researchers in Bangladesh returned to their three urban sites in 2005, five years after the original year-long survey, all three sites had been wholly or partly destroyed.

Fully addressing such vulnerability requires policymakers to use their powers to reduce insecurities rather than exacerbate them, and, with an active stance, to expand social security programs, improve health infrastructure and disaster management, and create stronger property rights. Collins et al (2009), for example, find that in South Africa health losses are substantially mitigated by the presence of free public clinics and the generosity of the pension system, based around a system of monthly government grants that gives households flexibility to cope with health losses (Case and Deaton, 1998). Implementing South African-style safety nets universally, however, would surely run up against budget and administrative constraints in most poor countries. While public sector solutions may be part of the picture, households’ main forms of risk coping will no doubt continue to be private.

It is not surprising then that providing low-income households with access to reliable and reasonably priced insurance mechanisms is increasingly taken to constitute a key part of “inclusive” financial sectors. The field of “micro-insurance” (a term that encompasses insurance products targeted to poor and low-income consumers) holds promise, but the field is young and no approaches have emerged so far that offer break-throughs akin to the original group-lending innovations that ignited the global explosion of microcredit (Morduch, 2006). Basic financial products like loans and savings accounts will remain critical devices for risk management in poor communities, just as they are in richer communities.

The persistence of shocks
The project of expanding insurance cover and coping mechanisms is made more urgent to the extent that temporary shocks translate into long-term losses. Collins, et al (2009) provide a string of stories of households hit by serious illness. Over time most of the households end up depleting their financial assets in their attempts to pay for medicines and doctors, and the households’ asset depletion translates into reduced earning potential over the long term. The most severe losses occur when the shocks themselves play out over time, as with a worsening case of tuberculosis that requires repeated visits to doctors and extended courses of medicine.\textsuperscript{13}

One sense of these dynamics emerges from papers that relate risk to health outcomes. Dercon (2004b) reports on a broad group of studies that link income shocks to health outcomes. Rose (1999), for example, correlates the incidence of bad rainfall realizations in rural India and increased infant mortality rates. She finds that the inability to cope with the temporary loss of income leads to choices that directly harm children, a finding that emerges in a sample of landless households (i.e., those with the most limited ability to self-insure) but not in the sample of households with substantial assets. The link between vulnerability to temporary income shocks and increased rates of child mortality is the most extreme example. Other examples, drawing on evidence mainly from poor populations in South Asia and sub-Saharan Africa, describe links between vulnerability to income swings and outcomes that fall short of death but which nevertheless generate lasting deprivations; they include low school attendance (Jacoby and Skoufias, 1997); increased child labor (Beegle, et al., 2003); reduced physical stature of children (Alderman, et al., 2002; Foster, 1995); and diminished school performance of young children (Alderman, et al., 2002; Foster, 1995). As with Rose’s (1999) study, these are not generic findings but hold chiefly for households with few assets, either of land or livestock, and limited means to self-insure.

The link between vulnerability and the profitability of enterprises is a staple of the theory of risk aversion (i.e., the notion that investors and entrepreneurs trade off average financial returns for a lower variance of expected returns), but it is not yet well-established in practice. The most promising avenue will likely involve studies that relate access to insurance to measures of productivity (e.g., the profitability of household business).

Dercon and Hoddinott (2004) sum up their evidence from Zimbabwe and Ethiopia by stressing the way that temporary set-backs turn into permanent deprivations: they find that the inability to cope with transitory shocks from droughts and other serious crises has long-term consequences, especially for children, for whom reductions in stature and

\textsuperscript{13} A sense of the magnitude of the problem is given by Angus Deaton’s (1992) simulations of optimal asset accumulation under borrowing constraints. Deaton begins with a dynamic stochastic choice problem in which a household builds up and draws down assets in order to dampen the variability of consumption in the face of income swings. Since by assumption the household cannot borrow, assets are used as buffers, and a relatively low level of assets can be used to smooth independent, identically distributed shocks. When the pattern or shocks has even a moderate auto-regressive component, a much higher average level of assets is required to smooth consumption.
schooling outcomes tend to diminish future employment prospects and productivity. Drawing on the empirical regularity that “taller (and better educated) women have, on average, taller (and healthier) children,” they conclude that “the impact of these transitory shocks may well be felt for several generations.” (Dercon and Hoddinott, 2004; p. 134).

Jalan and Ravallion (2004) tackle this question of transience and permanence of effects in structural estimation of income processes using a six-year panel (1985-90) of Chinese households in Guangdong, Guangxi, Guizhou, and Yunnan provinces. They investigate whether the error structure in income equations is consistent with poverty traps arising from temporary income shocks. They model the lagged dependent variable as a cubic function in a dynamic panel-data income model, looking for evidence of non-convexities in patterns of auto-correlated income. While income is found to be auto-correlated, such that a bad shock this year makes another bad shock more likely next year, Jalan and Ravallion do not find generalizable evidence that temporary shocks create poverty traps. They do, though, find large differences in the speed of recovery from shocks, with poor households taking far longer to bounce back than their better-off neighbors. The picture, as with the evidence above, is one in which risk is widespread and consequences are long-lasting for those who lack the means to cope.

7.1 Why insurance markets fail

The findings above help to show why risk management is important for poor households, yet the studies also points to a fundamental challenge: both the supply and demand of private insurance tend to be low in low-income communities, especially relative to the take-up of new credit and saving products.

The supply-side difficulties are stubborn, starting with the well-known problems of adverse selection and moral hazard (e.g., Besley, 1995). As the classic papers of Arrow (1963) and Pauly (1968) detail, information asymmetries can drive failures in the commercial provision of insurance. Systems of deductibles (households pay the cost of initial losses before insurers begin paying for the bill) and co-insurance (households cover a fraction of total expenses) can help, but practical difficulties remain when insuring outcomes that are heavily effort-dependent, such as crop failure and livestock health. When insurers cannot observe effort, nor observe the inherent riskiness of customers, contracts generally yield sub-optimal outcomes. In light of the theory, it is little surprise that there is so little profitable, large-scale commercial insurance coverage in low-income communities. Crop insurance accounts for a particularly notable gap given how important agriculture is in much of the developing world. Information problems make crop insurance a relatively less enticing product line for most commercial insurers (Morduch, 2006), but political imperatives make supporting farmers a priority policy initiative (Carter, et al., 2007). The combination leads to a landscape presently dominated by subsidy-dependent efforts.

In a theoretical treatment of the general insurance problem, Banerjee and Newman (1993) examine the market for risk from the viewpoint of the poor. While the Banerjee-Newman model focuses on broad relationships between risk-bearing and the income
distribution, an important insight emerges on the working of insurance markets. The logic of using deductibles and co-insurance to improve insurance markets hinges on exposing customers to enough risk such that they have incentives to work hard to limit bad outcomes. Providing complete coverage undermines such incentives. Banerjee and Newman (1993) show that in this set-up, it is the poor, rather than the rich, who – all else the same – will receive closer-to-complete insurance coverage. This is because poor households, who by definition live close to subsistence levels, are assumed to be relatively sensitive to variations in consumption levels when compared to richer households—an observation captured by the common assumption of declining absolute risk aversion (i.e., that as people get richer, they tolerate more risk). A given-sized deductible or co-insurance rate will thus deliver a greater dose of beneficial incentives for poor households than richer households. Equivalently, incentives can be delivered through the use of smaller deductibles and lower co-insurance levels when transacting with poor households than with rich households. Optimal contracts for poor households thus, in principle, deliver a greater extent of insurance cover relative to contracts for richer households.

Banerjee (2004) cites the literature on informal village insurance to argue that the poor may in fact be quite well insured. (For an overview, see Deaton, 1997; Morduch, 2006; Empirical studies include Townsend, 1994; 1995; Udry, 1994; Grimard, 1997; Fafchamps and Lund, 2003; Dubois, 2000; Jalan and Ravaillion, 1999; Ligon, et al., 2002; Morduch, 2004). But our review of the literature at this juncture suggests that the view is too optimistic, and that poor households remain substantially exposed to risk. Still, the polar extreme—in which household’s simply consume what they earn without smoothing ups and downs—is also a poor characterization. The literature so far shows that households do manage to self-insure and arrange collectively to share risk.

As Townsend (1994) suggests, there are several potential ways in which villagers might deal with risk on an informal basis, including:

1. **Diversification of a given farmer's landholdings into various spatially separated plots and into various crops,**
2. **Storage of grain from one year to the next,**
3. **Purchases and sales of assets such as bullocks and land,**
4. **Borrowing from village lenders or itinerant merchants and borrowing/lending more generally,**

In any given year villagers might well use more than one or even all of the mechanisms on this list. As Townsend notes, each of the mechanisms is by itself nontrivial to evaluate. Using survey data from high-risk villages in semi-arid India, Townsend (1994) looks holistically at the extent to which consumption among individual households covaries with average consumption in their villages. That is, through some combination of these or other mechanisms, do villages share risk? If markets for risk are complete, consumption should move together. Townsend finds that risk sharing, while not perfect, is quite good. In particular, credit and gifts are important in smoothing consumption; in some cases, the volume of loans and gifts exceeds average consumption (although
Morduch 1994 finds weaker evidence of extensive gift giving. Townsend goes on to look for evidence of regional risk-sharing but finds the data inconclusive.

Udry (1994) focuses in on credit as a source of consumption smoothing in the absence of insurance markets. By collecting a dataset of credit transactions in four villages in northern Nigeria, Udry provides insight into the specific nature of the mechanism behind the observed outcomes. The data are interesting even in their basic description: while over 75 percent of households lent money and nearly as many borrowed, the area was served neither by a formal financial institution nor specialized money lender. Nearly all loans (97 percent) were between neighbors and relatives. Loans were made without witnesses or even written records, and though the repayment amount was negotiated, explicit interest rates were never discussed. Only 3 percent of the loans were backed by collateral. Whether borrower or lender, 82 percent of those surveyed were able to enumerate the farm activities of the party on the other side of the transaction. It is this flow of information, Udry argues, that provides the basis for risk sharing. Most tellingly, payments are shown to respond to the financial circumstances of the lending household (through shorter payment periods and/or higher interest payments), a situation for which there is no provision in formal finance. (The reverse is seen as well: adverse shocks among borrowing households are met with lower payments over longer terms.) Nonetheless, Udry finds that the mutual insurance system provided by this arrangement is not sufficiently complete to insure all idiosyncratic risk faced by households.

Fafchamps and Lund (2003) build on Udry’s work by also considering gifts and transfers made outside of the context of loans, as well as savings and labor market participation. These steps have the benefit of increasing the scope for capturing transfers and responses to shocks, and they generalize findings to the village population, rather than just borrowers and lenders (though Udry finds they comprise the majority of the population in his sample). The authors’ panel data collected from the Philippines shows the majority (71 percent in value) of credit transactions are conducted between relatives and neighbors. More than 80 percent of loans are made within the same village, with virtually the rest occurring between neighboring villages. As in Udry’s data, lenders and borrowers are very familiar with each other’s activities: more than 85 percent of respondents were able to provide a complete accounting of the wealth holdings and demographic characteristics of their loan partners.

Gift giving is universal in the sample households: all households gave or received at least one gift, and 94 percent did so in each of the three survey rounds. 92 percent of households borrowed and 61 percent lent money, and 80 percent of informal loans carry no interest. For both gifts and informal loans the most common use of the funds is immediate consumption (a finding that squares with recent evidence collected by Johnston and Morduch, 2008 showing that on average half of microfinance loans issued to a sample of Bank Rakyat Indonesia customers were used for nonbusiness purposes). Loans respond to shocks (as in Udry, 1994), and gifts appear to as well, but with only borderline significance (p=0.13).
Fafchamps and Lund (2003) reject a village-level full insurance model in favor of one that works through networks of friends and relatives. Not all shocks are insured, and households respond to shocks in part by drawing upon financial savings (not livestock or crops), but not by increasing labor. As with Ligon, Thomas, and Worrall (2002, who revisit the data from Townsend 1994), Fafchamps and Lund find that the fit of their model is improved by taking account of limited commitment—i.e., through a model of insurance predicated on the notion that contracts in the informal sector are not enforceable and any party can withdraw from the scheme at any time.

Focusing specifically on illness, a major source of shocks for poor households, Gertler and Gruber (2002) find considerable exposure to risk in Indonesia. A key distinction drawn by Gertler and Gruber is the size of the shock: even if they take Townsend’s (1994) results as given and assume that minor health-related fluctuations in consumption can be smoothed over time, they argue that there may be less frequent, more serious health shocks that households will be unable to insure against through informal mechanisms. They test for this by looking at consumption patterns in a dataset incorporating a measure of the severity of illnesses. They find that while households are able to fully insure minor illnesses (those that do not limit physical functioning), they are only able to insure 71 percent of the economic costs (the cost of health care plus lost income) of moderate illnesses, and 38 percent of the cost of illnesses that severely limit physical functioning. The authors suggest the introduction of informal disability insurance as a potential remedy.  

Much of the literature above focuses on the village as a “natural” insurance unit, and the question posed by researchers centers on the degree to which villagers set up arrangements by which idiosyncratic risk is insured collectively. The structure of estimating equations captures the focus on idiosyncratic risk defined as income variability around the mean village income in a given year—i.e., doing better or worse than your neighbors. It is a helpful starting place, but the tests remain mute on the ability to cope with the kinds of major regional risks described in Dercon (2004): the broad devastation brought by wide-scale shocks like droughts, floods, and economic crises. In addition, villages are in many ways not natural insurance groups, especially relative to families and kinship groups that often extend geographically (Morduch, 2004; Munshi and Rosenzweig, 2007; Rosenzweig, 1988). Debate about the role of villages as risk-sharing collectivities goes back to Scott’s (1976) depiction of collectivist communities, contrasted with Popkin’s (1979) characterization of generally individualistic and inefficient relationships. The debate remains unresolved decades later, as evidence has accumulated on both sides. Goldstein et al (2004), for example, depict patterns of

---

14 One of the limits to this literature is given by difficulties in interpreting results. Consider the finding of Jalan and Ravallion (1999) on risk-sharing in China. They find that on average 40 percent of idiosyncratic income shocks translate into consumption shocks for the poorest households. Readers are left unclear as to whether the finding means that all poor households are vulnerable to 40% of shocks—or might, say, half fully insure while the other half suffers 80 percent of the loss? Moreover, are positive shocks handled differently from negative shocks? These kinds of questions are less important if the focus is on testing the basic fit of a benchmark model of full consumption insurance as in Townsend (1994), but they are critical in evaluating and developing policy responses.
inclusion and exclusion in community-level risk-sharing arrangements in Southern Ghana, drawing on insights into incomplete informal insurance (for more on fragmented collective relationships, see Fafchamps and Lund, 2004; Genicot and Ray, 2003; Platteau, 2000). This strand of literature shows how and why households may remain vulnerable even to idiosyncratic risks, in addition to broad aggregate risks. Grimard (Grimard, 1997) looks for evidence of risk sharing across regions in Côte d’Ivoire. He, too, rejects complete insurance but finds evidence of partial insurance between members of the same ethnic groups, especially in regions with the lowest availability of formal financial products. As Grimard suggests, the lack of full regional insurance might not come as a surprise given the difficulty in monitoring and enforcement over long distances. Munshi and Rosenzweig (2007) provide complementary data on intra-caste and intra-family insurance in India.

Taking this literature as a whole, it becomes clear that insurance provided by the formal sector should be seen in the context of a broader array of risk-coping mechanisms employed by households, some deployed preventatively and some used after shocks have occurred. Formal sector insurance includes a “loading factor” that includes taxes and administrative costs, and, as a starting point, insurance will be demanded only if compared to other options it is both relatively effective and relatively cheap.

7.2 Partnership models and index-based insurance

 Taken as a whole, the empirical literature on informal collective insurance shows that poor households remain substantially without insurance (especially when aggregate risk is considered) while richer households tend to be better insured. The literature on information asymmetries suggests that the gaps in coverage are not just inequitable but apt to be inefficient as well.

The literature on informal credit and saving has offered guidance for developing commercially implementable credit and saving products, but parallels are harder to find with regard to insurance. One hurdle for commercial insurers is posed by the logistical challenge of collecting small-sized premia from customers and needing to assess claims for losses that may loom large for small households but which are relatively small for major insurers (Morduch, 2006). The pursuit of profitability makes serving the poor with current technologies a low commercial priority.

Two new approaches are starting to change the equation. The first is partnership models, whereby commercial insurers partner with microfinance institutions to deliver commercially viable products. The insurers bear the major risks and maintain responsibility for actuarial calculations and pricing strategies, while the microfinance institutions use their existing relationships with customers to sell products and handle claims. The model has been especially successful in delivering “credit-life” insurance, in which microfinance customers receive term life insurance coverage during the life of their loans. The product is relatively simple to administer, and premia are typically collected as extra fees on top of the interest rates paid for loans. A disadvantage is that
this pricing may be opaque to borrowers and many products appear to be costly relative to the value of coverage provided. Also coverage extends to borrowers only—coverage lapses if customers choose to stop borrowing. Still, the principle is robust, and the promise remains for developing an improved array of insurance products.

The second new approach is index-based insurance (Carter, et al., 2007; Skees, et al., 2004). As high transaction costs, moral hazard, and adverse selection have thwarted attempts to provide crop insurance on a commercial basis and wide scale, new approaches have sought to side-step those problems by shifting from insuring crop losses to insuring bad weather realizations instead. The insight is that if the correlation between crop losses and bad weather is high enough, substantial insurance can be provided through index-based weather insurance. Farmers are powerless to change the weather; thus moral hazard and adverse selection disappear. Transactions costs also fall since claims do not need to be verified and products can be standardized around a given weather station.

An example is given by a rainfall insurance product offered in Andhra Pradesh, South India studied by Giné, Townsend, and Vickrey (2007b). To get a sense of the product, it is worth reviewing contract details. The insurance contract divides the cropping season into three parts, roughly corresponding to sowing, podding/flowering, and harvest phases, and farmers can purchase separate contracts for each part. The risk in the early phases is that rainfall will be insufficient, so, in this example, the contract pays nothing if rainfall exceeds 70 mm. If accumulated rainfall is less than 70 mm, the policy pays 10 rupees for each millimeter of rainfall below the cutoff, paying out a fixed amount (1000 rupees) when the season is extremely dry. In the third (harvest) phase, problems emerge when rainfall is excessive, so the policy reverses itself; it now pays out when rainfall exceeds 70 mm and pays nothing below the threshold. A policy covering all three phases is inexpensive enough to be accessible to low-income farmers (coverage costs 200-300 rupees or US$5-6; Giné, et al., 2007a).

In principle, even villagers who are not farmers can purchase contracts. While crop insurance is marketed only to farmers, there is nothing stopping the sale of weather insurance to anyone in the region who wants protection from the ups and downs of weather-related demand and supply fluctuations. Since the risks are correlated locally, prudence requires that rainfall insurance be offered in partnership with a re-insurer who can help local retailers spread risks across regions. The large global re-insurers like Munich Re and Swiss Re are playing a role in spreading risks across countries and broad regions.

**Low demand**

The idea of rainfall insurance makes sense. As a reality, though, the product in Andhra Pradesh described above has not been embraced enthusiastically by farmers, despite its relatively low cost and the major costs of drought in the region (Giné, et al., 2007a). The reasons why most people refused to buy it are hard to pin down, but basis risk between insurance payouts and the risk insured is a major determinant.
The value of rainfall insurance depends on a high correlation between incomes and rainfall as measured at the local rain gauge or weather station; the divergence is so-called basis risk. Two forces combine to create basis risk. First, there may be an insufficient number of rainfall gauges to adequately capture weather variations in a region. Second, the impact of weather on a given plot may be affected by the characteristics of the plot, including its slope, soil quality, and the availability of alternative water sources; the contract’s pricing structure thus may not make sense for all farmers equally, even when they face the same weather patterns. Not surprisingly, Giné et al. (2007a) find that take-up of the Andhra Pradesh rainfall insurance product decreases with basis risk.

They also find that demand increases with household wealth and decreases with the extent to which credit constraints bind; again neither is surprising although since it is the poorest households that are generally least able to bear risk, the opposite finding (that take-up decreases with wealth) would also be plausible. One clearly surprising result from Giné et al. (2007a) is that take-up falls with risk aversion. Since those who exhibit the greatest degree of risk aversion should be most eager to obtain insurance, the result falls outside the benchmark model. The most likely explanation is that it is uncertainty about the product itself (Is it reliable? How fast are pay-outs? How great is basis risk?) that drives down demand.

The evidence on low demand is repeated with a similar product in a very different setting. Giné and Yang (2008) investigate the demand for a hybrid credit-rainfall insurance product in Malawi. Their hypothesis is that risk-averse farmers will be unwilling to adopt new agricultural technologies—in this case, high-yielding varieties of maize and improved groundnut seeds. Coupling insurance with a credit product should then, in principle, increase adoption. The study employs a randomized field experiment involving about 800 maize and groundnut farmers; half the farmers were offered credit to buy the new seeds, while the other half were offered a similar credit product coupled with a weather insurance policy. The insurance policy was priced at actuarially fair rates, providing some protection against low rainfall outcomes. (Though, we note that given that there is basis risk, a rate that is actuarially fair in terms of weather realizations at the rainfall gauge may not be actuarially fair on a given farmer’s plot.) A third of the farmers who were offered just the credit contract accepted it. Surprisingly, take-up was 13 percentages points lower for the half of the sample offered the combined credit-insurance product. One possible explanation discussed in the working paper version is cognitive: insurance is not a simple concept and the terms of the policy can be hard to weigh, a contention supported by the finding that more educated farmers were more likely to take up the combined product in the Malawi experiment. Increased exposure to the product, financial literacy training, or a new marketing strategy may all play a role in raising demand levels. Another explanation is that farmers already received insurance implicitly through loan contracts: in case of a serious drought, farmers do not have to repay loans (by virtue of limited liability); this, in itself can limit demand for the insurance product. As with the case of the Indian product, it is also critical that the product itself be desirable in more basic ways (with reasonable administrative costs, reliably speedy pay-outs, and a sufficient spread of rainfall gauges to ensure limited basis risk--factors that are not
always in place). In the scale-up of this product, the firms in Malawi no longer sell directly to farmers, and are instead insuring the lenders and large firms directly.

Carter et al. (2007) argue that a variation on index-based agricultural insurance can sometimes do better. This is area-based yield insurance, in which payouts are based on measured average yields in a region (rather than on yields on the policymakers own plot). Basing pay-outs on average yields in, say, a valley, effectively eliminates moral hazard and adverse selection in the same way that weather-based insurance does (Miranda, 1991). The advantage is that coverage can extend beyond risks due to weather only--a finding stressed in the Carter et al. (2007) application in northern Peru. A disadvantage is that, as with traditional crop insurance, yields need to be measured, a sometimes expensive and time-consuming process. And, as with rainfall insurance, basis risk remains a problem for customers.

These studies provide a start at unpacking the reasons that the demand for insurance tends to be low. Additional reasons include the fact that the effectiveness of informal insurance mechanisms may, in some cases, be sufficient to limit the net impacts of formal insurance contracts (Morduch, 1999), and the expectation that in severe disasters the public sector will step in with aid, again limiting the net impact of private insurance. The importance of marketing in determining demand for credit (Bertrand, et al., 2010) also highlights the urgency for insurers to identify new modes of marketing to enhance take-up of insurance. Given that the prevalence of risk is widely taken to be a fundamental element in the lives of poor households, it is notable that such fundamental questions about the demand for insurance remain; the topic is a priority on the research agenda.

7.3 Health Insurance

Health risks loom large among the risks faced by poor households, and out-of-pocket health spending is high. In the United States, the percentage of health spending that is out-of-pocket (i.e., not paid by insurance) is about 12 percent (Pauly, et al., 2008). In Bangladesh, the comparable figure is 64 percent; in Ghana, 59 percent; in India, 78 percent; in Paraguay, 55 percent; in the Philippines, 47 percent; and in Vietnam, 62 percent. The figures are from the 2002 World Health Survey, a World Health Organization household survey covering about 4,000 to 6,000 households per country (cited in Pauly et al, Table 1).15

As with crop insurance, few commercially successful health insurance programs exist that serve poor communities (Morduch, 2006). And, as with the other types of insurance discussed above, moral hazard and adverse selection pose substantial barriers for insurers. In line with theory, co-payments can help. Grameen Kalyan, a Bangladesh-based health insurance scheme that is part of Grameen Bank offers coverage for preventative and curative health services, and requires co-payments for the curative services. Co-payments turned out to help Grameen Kalyan not only reduce over-use of

\[15\] Most of the spending is attributable to spending on prescription drugs. Turning to the same countries, the percentages are: Bangladesh, 84 percent; Ghana, 48 percent; India, 55 percent; Paraguay, 73 percent; the Philippines, 61 percent; and in Vietnam, 44 percent (World Health Survey, cited in Pauly et al, Table 3).
medical services, but they helped to effectively signal the quality of care as well (given that customers judged quality by price; Radermacher et al., 2006, p. 78). Elsewhere, though, co-payments have been viewed skeptically in poor populations. In Mali, for example, co-payments were feared to be too onerous for poor customers, and a system of cooperatives was founded to help members pay co-payments for health services (Radermacher, et al., 2006, pp. 78). It is unclear whether the problem here is with requiring co-payments *per se* or with the levels at which they are set—a problem ripe for research.

One of the themes of the chapter is that design matters, and the sentiment is echoed here with regard to health insurance. Above all else, insurers sell their reputations: the promise companies will reliably and promptly deliver pay-outs when crises hit (in the exact amounts and with the exact timing specified in contracts). Such basic reliability cannot be taken for granted in either the public or private sectors (Das, et al., 2008). If doctors, nurses, and pharmacists are unavailable or of uncertain quality, a reliable health insurance system is hard to envision. Yet experiences with hospitalization insurance in South India suggest that this conclusion may put things backward. In the health insurance program of BASIX, for example, the organization of villagers into insurance groups, facilitated by their membership in BASIX’s microfinance program, created a large enough block of customers that it was possible for the insurer to certify and contract with high-quality doctors directly. This “preferred provider” system benefited doctors and medical personnel who gained from the steady demand. It was thus the existence of the health insurance program, and the financial resources it aggregated, that helped fix quality deficiencies in healthcare quality faced by customers. In other programs, like that of BRAC and Grameen Kalyan in Bangladesh, nearly all health care is provided within the programs’ own clinics (Radermacher et al., 2006, pp. 86, 91).

Another concern rests with cost. Without reliable data, insurers face a series of risks in pricing insurance, most importantly, mis-estimating the probabilities of loss and fluctuations in health care costs (Radermacher et al., 2006, p. 90). But even if it is possible to set rates precisely, it is unclear that customers are willing to shoulder the full costs. Product design features appear to help. Some institutions lend money to customers to help them pay premiums: FINCA Uganda, for example, and Karuna Trust in Karnataka India. Others break the premiums into small-sized installments that can be easier for poor households to handle (though when given a choice, BRAC customers sought monthly rather than weekly installments). But most of the health insurance programs assessed by Radermacher et al. (2006) nevertheless lost money. The unanswered question is whether, if offered a truly effective high quality health insurance policy, households would willingly pay the required price. As it is, many customers fail to renew their contracts at the end of a given year—Radermacher et al. (2006, Table 10) show BRAC’s renewal rate at 51 percent, Grameen Kalyan’s at 54 percent, and VimoSEWA’s (part of a long-established community institution in Ahmedabad, India) at 51 percent. Better marketing and rising comfort levels will likely help raise renewal rates, but the evidence on renewal rates suggests a *prima facie* case that success will also require improving service quality for the price. This is an area that will surely benefit
from the kinds of experimental economic research that has energized the literatures on credit and saving.

8. Governments, businesses, and non-profit institutions
Muhammad Yunus (Yunus, 2008), in defense of his strategy in founding Grameen Bank, argues for expanding financial access by launching “social businesses.” In this case, this means banks that earn modest profits and that do not distribute them to investors; instead profits are re-invested in institutions to further social goals. Such social businesses are necessarily dependent on donors and other “social investors” who value the institutions’ social aims and who are willing to sacrifice financial returns in order to obtain greater social returns. Economists in the tradition of Milton Friedman (1970) argue instead that the optimal path entails maximizing pure profit. Surpluses can then be distributed to improve social welfare. In this vision, there is no role for social businesses.

Resolving these conflicting visions has been hampered by the lack of evidence on the impacts of subsidy, allowing ideological positions (on all sides) to go unchecked. The specter of inefficient, heavily subsidized public banks hangs over the microfinance movement, and it is partly responsible for a lack of enthusiasm for embracing donated funds. Microfinance advocates, wary of subsidies, have energetically sought to move microfinance to a commercial basis, free of long-term subsidies. Their goal is to expand scale while not sacrificing mission. There is thus a general consensus among microfinance leaders, that problems tend to arise when the government becomes a direct lender (United Nations, 2008); the main problems include inefficiency and the mistargeting of subsidy.

Conning and Morduch (2007), though, argue that while commercial investors play important roles, by virtue of their philanthropic objectives, social investors can improve social welfare and total surplus in ways that commercial investors by themselves often cannot. Their theoretical model is built around a nested moral hazard problem, and draws on approaches to corporate finance (e.g., Dewatripont and Tirole, 1994; Tirole, 2005). At one level is the well-known moral hazard problem in credit markets, in which borrowers lack collateral and thus contracts are maximized subject to a limited-liability constraint. In these second-best contracts, optimal levels of effort are generally impossible to elicit (i.e., effort to ensure that risky ventures work out well). A second layer of moral hazard exists between investors and institutions (Hansmann 1996). Just as lending to villagers entails risk, investing in micro-lenders also entails risk, and such risk is also not generally backed by collateral.

In this setting, Conning and Morduch (2007) show cases in which philanthropic giving can raise output. Thus, not only does it reduce poverty but it can also increase an economy’s overall productivity. The possibility arises because social investors, by definition, are willing to absorb costs in order to bring gains. (This is, of course, their reason for existence.) They can then do more than shift resources: they can also help increase the total sum of resources available.
The result emerges when social investors expand the contract space by being willing to subsidize the “limited-liability rents” that often stand in the way of trades in the commercial loan market. (Limited-liability rents refer to the returns that must be left with borrowers, or the loan officers that monitor them, in order to provide incentives to work hard even when limited liability constraints restrict incentives created by threatened punishments.) When such rents become large relative to expected project gains, lenders may opt not to make the loans, leaving socially valuable projects unfunded. A small subsidy in such contexts may be enough to make lending profitable. If enough small loans become profitable, a new microfinance institution might emerge where otherwise it might not have.

In the Conning-Morduch (2007) framework, profitability is neither necessary nor sufficient for attracting commercial capital given that risky investments remain unsecured. The framework implies that the most effective uses of social investment are either “high” or “low” when directed to poor communities. In the “high” case, optimal subsidy involves allocations that either go to institutions serving relatively well-off (but still poor) customers, where some collateral is obtainable and the potential for leverage is maximized (i.e., commercial funds can be drawn in). In the “low” case subsidies go to institutions serving the poorest customers where social gains are greatest. There is less “bang for the buck” in supporting a middle range of institutions that are self-sufficient with their own resources but for which leverage (and potential scale) is limited. The Conning-Morduch framework thus embraces a range of strategies to promote financial access.

In practice, though, the microfinance industry has divided around competing strategies. Cull et al. (2009b) shed light on the trade-offs between those competing strategies. They employ a data set that includes 346 of the world’s leading microfinance organizations and covers nearly 18 million borrowers. One strength of the data set is that the data are adjusted to show the roles of both explicit and implicit subsidies (generally delivered as soft loans). The data suggest that, while commercialization is a powerful trend, commercial banks and those that are avowedly “social businesses” like Grameen Bank are not substitutes. The data show that microfinance is not taking a single path, nor that it should. The authors put forward a series of basic findings to frame debates:

1. Licensed commercial banks with for-profit status serve a growing share of low-income customers and garner media attention, but the share remains relatively small. Globally, microfinance continues to be dominated by nongovernmental organizations (NGOs), government-owned banks, and “non-bank financial institutions” that are a cross between banks and NGOs.

2. Being a non-profit institution does not mean being unprofitable. A large share of microfinance institutions with “non-profit” status in fact earn steady profits—but

---

Cull et al (2008) report that the adjustments include “an inflation adjustment, a reclassification of some long-term liabilities as equity, an adjustment for the cost of subsidized funding, an adjustment for current-year cash donations to cover operating expenses, an in-kind subsidy adjustment for donated goods and services, loan loss reserve and provisioning adjustments, some adjustments for write-offs, and the reversal of any interest income accrued on non-performing loans.”
(as “social businesses”) they re-invest their profits in the institution and cannot legally distribute earnings to shareholders. Earning profits (and thus limiting dependence on subsidies) and becoming a commercial entity are distinct activities. Neither implies the other.

3. Commercial microfinance banks, as a group, make loans that on average are about four times larger than loans from NGOs. Since poorer customers generally demand smaller loans, average loan size is a rough proxy for the poverty level of customers. On average, commercially oriented microfinance banks thus tend to serve a substantially better-off group of borrowers than do NGOs. These microfinance banks, as a group, also serve fewer women as a share of their customers.

4. Most microfinance institutions charge inflation-adjusted interest rates between 20 and 40 percent per year. NGOs as a group charge interest rates that are roughly double the size of the average charged by commercial microfinance banks. Thus, the poorest customers tend to pay the highest interest rates on loans. The high interest rates are necessary to cover the added costs of making small loans, and the NGOs’ record of expansion and high loan repayment rates over time suggests that customers value the services, even at high costs.

5. Despite the high interest rates, most of the institutions serving the poorest customers earn profits too small to attract profit-maximizing investors.

6. Thus, subsidies and noncommercial funding continue to be important to nongovernmental organizations, while banks rely mainly on social investment and commercial sources of capital.

The findings are generally in line with Conning’s and Morduch’s framework, which suggests that debating about a single, correct vision for microfinance—be it a non-profit approach or a commercial model—misses the reality that microfinance flourishes thanks to a diversity of strategies. One key question, that remains unanswered, hinges on whether subsidized microfinance institutions serving the poor are likely to unfairly compete against commercial institutions. The Cull et al. (2009b) evidence suggests that this is unlikely since being subsidized does not mean being inexpensive; as noted above, even with subsidies, poor households pay more for credit than richer households.

The role for governments remains unclear. While state-run banks have a generally poor record as lenders, government savings banks and postal savings schemes have proven successful around the world. Beyond that direct role, policymakers find an important role in setting appropriate regulation and creating a supportive environment for expanding financial access. As Ahlin and Lin (2006) suggest in cross-institution regressions, macroeconomic environments featuring steady growth and low inflation not surprisingly improve the performance of microfinance institutions.

Cull et al. (2009a) argue that regulation is critical, especially for deposit-taking microfinance banks, but they demonstrate that regulation carries costs. They draw on a database that combines financial data on 245 microfinance institutions with data on levels of prudential supervision. Their basic regressions show that regulatory supervision is negatively associated with profitability. Once the non-random assignment of supervision
is addressed via treatment effects and instrumental variables regressions, supervision appears to be associated with substantially larger average loan sizes and less lending to women than in the least squares regressions, though it is not significantly associated with profitability. Cull et al. (2009a) conclude that, as a group, profit-oriented microfinance institutions absorb the cost of supervision by curtailing outreach to market segments that tend to be more costly per dollar lent. Cull et al. (2009a) do not argue against supervision, but argue that policymakers should take trade-offs into account.

9. Concluding comments
In the past decade, research on financial access for the poor has made a transition that parallels the success of practitioners in bringing successful microfinance programs to communities around the world. The earlier mode for research on financial access centered on understanding the financing constraints and opportunities of poor households. This work laid out the case for innovations, and researchers aimed to quantify the potential gains from relaxing financing constraints. For the most part, the scenarios were hypothetical given that actual banking institutions capable of providing reliable, professional financial services in poor communities were scarce. In focusing on understanding household needs, research on financial access had much in common with research on healthcare and education in poor parts of developing countries. Researchers provided insight into households’ coping mechanisms and helped draw the outline of optimal solutions.

In the past decade, the financial landscape has transformed in villages and slums worldwide, and so have important strands of scholarship. Gaps in financial access remain wide, but viable institutions dedicated to serving poor and low-income communities have penetrated thickly in Bangladeshi villages and are making substantial in-roads in rural India. In Latin American and Africa, most of the action has been urban so far, but innovation has been rapid. The spread of mobile telephones, for example, heralds the possibility of expanding financial services widely via mobile banking in sub-Saharan Africa. Customers in parts of the Philippines, Peru, and Bolivia can now choose between a range of different microfinance providers actively seeking their business.

This chapter illustrates how much is being learned from this expansion of activity—and how much remains to learn. If there is a single message of the chapter, it is that mechanisms matter. The design of products (including their prices, term structure, flexibility, and marketing) affects adoption and usage—and ultimately economic and social impacts. That insight is made visible in part through a shift in research strategies. Much of the new work described above involves researchers creating partnerships with existing financial providers, both commercial banks and public or non-profit institutions. The partnerships generally center on testing the effects of systematically varying delivery mechanisms or introducing new services. The lessons are thus not hypothetical but are based on actual products delivered by actual institutions. Most of the lessons are, by definition, directly implementable.

This shift in research strategy puts a natural focus on innovation, with practitioners seeking to expand existing approaches or test competing strategies. The practice of
medical and consumer product and marketing trials has given the new research much of its direction, including the focus on randomized controlled trials. While trial-based approaches are not the only way that researchers are analyzing trends and possibilities, they have become an important part of the mixture of methodologies. The trial-based methods have been particularly valuable in testing long-held assumptions and opening the door to behavioral perspectives that introduce elements of psychology into the economic framework. One important lesson that the trials yield, and one which applies as well to sectors like health and education, is that modes of implementation matter. The new research orientation complements the traditional more observational approach to empirical economic research by working with functioning institutions, identifying needs that they might fill through innovation and adaptation, and designing studies to test theories fundamental to individual decision-making and functioning of markets.

References


212. Tripati, S. (2006). Microcredit won't make poverty history (Small loans can empower women, argues Salil Tripathi - but don't expect them to help the very poorest, no matter what the Nobel judges say), Guardian Unlimited (online).


