# The Basic Analytics of Access to Financial Services

**Thorsten Beck and Augusto de la Torre***

**First draft: May 2005**

**This draft: March 2006**

**Abstract:** Access to financial services, or rather the lack thereof, is often indiscriminately decried as problem in many developing countries. This paper argues that the "problem of access" should rather be analyzed by identifying different demand and supply constraints. We use the concept of an access possibilities frontier, drawn for a given set of state variables, to distinguish between cases where a financial system settles below the constrained optimum, cases where this constrained optimum is too low, and—in credit services—cases where the observed outcome is excessively high. We distinguish between payment and savings services and fixed intermediation costs, on the one hand, and lending services and different sources of credit risk, on the other hand. We include both supply and demand side frictions that can lead to lower access. The analysis helps identify bankable and banked population, the binding constraint to close the gap between the two, and policies to prudently expand the bankable population. This new conceptual framework can inform the debate on adequate policies to expand access to financial services and can serve as basis for an informed measurement of access.

**JEL classification codes:** G18 ; G21 ; G28 ; O16

**Keywords:** Access to Financial Services; Credit Market Imperfections; Competition; Regulatory Policies

---

* Thorsten Beck is a senior economist in the Development Research Group of the World Bank; Augusto de la Torre is the Regional Financial Sector Advisor for Latin America and the Caribbean of the World Bank. This paper has been prepared as part of the Latin American Regional Study on Access to Finance. It has benefited from discussions and comments by Abayomi Alawode, Ole Andreassen, Jerry Caprio, Stijn Claessens, Asli Demirguc-Kunt, Patrick Honohan, Martin Jung, Soledad Martinez Peria, Sergio Schmukler, Tova Solo, Robert Stone, as well as from comments and suggestions from participants at various seminars in the World Bank, the Asobancaria seminar on Bancarización in Cartagena 2004, the World Savings Banks Institute conference on Access to Finance in Brussels 2004 and the Dia de la Competencia in Mexico City 2005. This paper's findings, interpretations, and conclusions are entirely those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

# I. Introduction

Access to financial services or outreach of the financial system has become a major concern for many policymakers in developing countries. While the use of financial services—measured as having deposit accounts with banks—reaches over 90% in most high-income countries, in many low- and even middle-income countries the use of formal financial services is still restricted to a small number of firms and households (Peachey and Roe, 2004; Beck, Demirguc-Kunt and Martinez Peria, 2005). Moreover, the intense financial sector reforms undertaken by many emerging economies over the past decade—doing away with interest rate controls and directed credit, liberalizing entry and privatizing state-owned banks—have not led to the type of broadening of access to financial services that was initially expected, particularly for lower-income households and small and medium-size enterprises (SMEs).

Broad access to financial services is related to the economic and social development agenda for at least two reasons. First, a large theoretical and empirical literature has shown the importance of a well developed financial system for economic development and poverty alleviation (Beck, Levine and Loayza, 2000; Beck, Demirguc-Kunt and Levine, 2004; Honohan, 2004a). To be sure, while a causal link running from financial depth to growth has been rather convincingly established by empirical research, the search for causality between the breadth of access and growth is still on. However, as noted by De la Torre and Schmukler (2006a), the discussion of the plausible channels through which financial depth could cause economic growth often resorts to access-related stories. Prominent in this regard is the Schumpeterian view that finance leads to growth because it fuels "creative destruction" by allocating resources to efficient newcomers. That is, through broader access to external funds, talented newcomers are empowered and freed from the disadvantages that would otherwise arise from their lack of

inherited wealth and absence of connections to the network of well-off incumbents (Rajan and Zingales, 2003). Second, access to financial services can be seen as a public good that is essential to enable participation in the benefits of a modern, market-based economy, in an analogous way as is the access to safe water, basic health services, and primary education (Peachey and Roe, 2004).

A low level of observed use of financial services, however, has to be carefully distinguished from a problem of access. In a purely theoretical—and rather uninteresting—world characterized by the absence of transaction costs, uncertainty, and asymmetric information there is no "problem" of access. Decisions to accumulate savings, take out loans, and make payments would be equally open to all and the implementation costless. Banks would not be needed to mobilize savings, facilitate payments, and allocate loans, as savers would assign their savings directly to borrowers based on perfect knowledge of investment possibilities. Hedging or insurance products would not be required given the absence of uncertainty. Access to external finance would be frictionless, limited only by the inter-temporal wealth constraint of the borrower, which would be known equally well and with certainty by both the lender (saver) and the borrower (investor). Investment decisions would be independent of financing and consumption decisions. The choice between borrowing and lending (saving) would be determined purely by inter-temporal preferences and investment opportunities, and changes in borrowing and lending would reflect changes in demand and investment opportunities rather than changes in the possibility of access. In this ideal world, the lack of use of finance by some agents would certainly not be a "problem" in the commonly used sense of the word. Agents that do not borrow for consumption would be those that do not need to smooth their consumption

over time subject to their life-time wealth. And projects that do not borrow for investment would be those that do not meet the condition of a positive real net present value.

Problems of access do arise in some well-defined sense, however, in the real world and are essentially linked to such crucial real-world facts as transaction costs, uncertainty about project outcomes, and information asymmetries. These introduce frictions that can limit access to financial services and that can make it difficult to de-couple investment from financing decisions in most cases. In a world with frictions, investment decisions may reflect credit supply constraints, and not just preferences and business opportunities. It is precisely these frictions which give rise to organized financial markets, financial institutions, and broader contractual entities.[1] However, the efficiency with which financial markets and institutions overcome these frictions depends on the macroeconomic environment, market structure, and overall contractual and informational environment. Across countries and over time, we can observe a large variation in the efficiency with which financial markets and institutions are able to overcome market frictions and provide financial services.

To say that problems of access to financial services arise due to transactions costs, information asymmetries, and uncertainty does not entail, however, that an access "problem" is always easy to identify. On the demand side, economic agents, households and enterprises alike, might have no impediment to access financial services, but may simply not want to use them. It would be wrong to argue that voluntary self-exclusion constitutes a "problem of access," except in the cases where self-exclusion reflects unduly low levels of financial literacy or is a

[1] Transactions costs, agency problems, and uncertainty are key reasons why institutions and organizations exist (North 1990). In a world with financial market frictions, basic financial services are typically categorized into savings, loan, insurance, and payment services. By offering payment services, financial institutions and markets reduce transaction costs in the exchange of goods and services between people and over time. By offering savings and loan services financial institutions and markets allow firms and households to overcome frictions that prevent them from de-linking consumption from investment decisions as discussed above. By offering insurance mechanisms, financial intermediaries allow households and firms to hedge and diversify risks and smooth consumption. Compare the overview in Levine (1997, 2005).

psychological response to past systematic discrimination. On the supply side, creditors that face large macroeconomic risks and/or major difficulties in mitigating problems of adverse selection, moral hazard, and contract enforcement may decide to deny loans to certain borrowers. Doing so would be a matter of prudence in the use of depositors' funds. Again, whether this situation constitutes a "problem of access" is debatable and the opposite is easier to argue—that it would be a serious "problem" if creditors made loans to certain borrowers under such circumstances, as many a banking crisis illustrates. The key point is that, once the existing market frictions in an economy are taken into account, the observed lack of use and outreach of financial services might be the rational and prudent outcome. But, should such an outcome deserve the label of a "problem"? And if so, in what sense?

Traditionally, access problems have been defined by reference to some form of observable limitation that leads to a contrast between the active use of a given financial service (say, a loan) by a certain group, on the one hand, and the low use (or lack of use) of that service by another group, on the other hand. Thus, we talk about *geographic limitations*—reflected, for instance, in the absence of bank branches or delivery points in remote and sparsely populated rural areas that are costlier to service.[2] We also talk about *socio-economic limitations*—when financial services appear inaccessible to specific income, social or ethnic groups either because of high costs, rationing, financial illiteracy, or discrimination. And we also talk about *limitations of opportunity*—where, for instance, talented newcomers with profitable projects are denied finance because they lack fixed collateral or are not well connected.[3]

[2] Beck, Demiguc-Kunt and Martinez Peria (2005) find a positive cross-country association of geographic branch and ATM penetration with population density and physical infrastructure.

[3] In an alternative classification Honohan (2004b) distinguishes between price barrier (a financial service is available but too expensive), information barrier (a firm's or household's credit worthiness cannot be established) and product and service barrier (services most needed by certain groups are not offered).

While limitations to access along these three dimensions are due to market frictions, the observed outcome could be a constrained optimum—the result of rational agents maximizing their utility and profit functions given the constraints imposed by the existing market frictions. In what cases and in what sense, then, can we say that a low level and unevenly distributed access constitute a "problem"? Our approach in this paper is to define the "problem of access" in terms of an "access possibilities frontier," which is drawn for a given set of "state variables." In our framework, thus, an access problem is defined in three different ways: (i) when the economy settles at a point below the access possibilities frontier, given the state variables; (ii) when the possibilities frontier is too low relative to countries with comparable levels of economic development; and (iii) when imprudent lending practices lead to an excessive credit expansion beyond the constrained optimum.

To simplify the discussion we will bundle the market frictions that are relevant to the supply of financial services into two groups: (i) transaction costs and the resulting scale economies of financial services at the level of the user, the institution, and the market, and (ii) systemic and idiosyncratic risks. On the demand side, we will differentiate between economic and non-economic factors that may lead to self-exclusion. While this is clearly a major simplification of the access problem, it will help us derive an analytical tool to better discuss access issues and relevant policies. In section 2, we analyze access to simple payment and savings services. In section 3, we examine access to lending services. Section 4 concludes.

## II. Analytics of Access to Payment and Savings Services

This section discusses supply and demand factors for access to payments and savings services. While there is a considerable diversity of payments and savings services, even in relatively underdeveloped financial systems, our analysis focuses on the demand and supply of the most plain-vanilla versions: a payment service based on a simple checking account and instrumented through either a check or a debit card, and a saving service consisting of a passbook savings account that pays a zero real interest rate and is redeemable at par and on demand.[4] We, therefore, emphasize in this section the transactional and custodial functions of these services, respectively, rather than their interest-earning dimension. We further assume that deposits in checking and passbook savings accounts are invested by banks in risk-free securities (a narrow bank scenario). Hence, for the purposes of our analysis, the price of payments and savings services is given by a fee, the intermediation margin earned by banks on the corresponding accounts is negligible, and risk considerations are of no relevance.[5] These simplifying assumptions—which imply no loss of generality for the analysis of issues in access to this type of services—allow us to focus on costs as the driver behind the supply of payments and savings services. Although we consider only plain-vanilla fee-based services, we do make a distinction that matters for access, as discussed below—the distinction between the production of high-value payments and savings services, on the one hand, and low-value/high-volume services, on the other.

[4] We assume, therefore, that checking and savings accounts pay a nominal interest rate that is equal to the local inflation rate.

[5] To be sure, even in the narrow-bank scenario assumed here, there are some forms of risk, notably operational risk. Clearly, risks (credit, liquidity, price, etc.) come into the center of the stage as we depart from the narrow-bank assumption and banks invest the sight deposits into loans and other risky assets. We introduce credit risk as a core element of the analysis of access to lending services in the next section.

### *(a) Fixed transaction costs*

Fixed transaction costs in financial service provision result in decreasing unit costs as the number or size of transactions increase. These fixed costs exist at the transaction, client, institution, and even financial system level. Processing an individual payment or savings transaction entails costs that are, at least in part, independent of the value of the transaction. Maintaining an account for an individual client also implies costs that are largely independent of the number and size of transactions the client makes. At the level of a financial institution, fixed costs are crucial and span across a wide range—from the brick-and-mortar branch network, to computer systems, to legal services, to accounting systems, and to security arrangements—and are rather independent of the number of clients served or the number of transactions processed. Fixed costs also arise at the level of the financial system, including in terms of regulatory costs and the costs of payment clearing and settlement infrastructure, which are again, and up to a point, independent of the number of institutions regulated or participating in the payment system. The resulting economies of scale at all levels make it unprofitable to stay in the business of payment and savings service provision unless the associated scale economies are captured in some form.

The effect of fixed costs on financial service provision can be reinforced by network externalities, where the marginal benefit to an additional customer is determined by the number of customers already using the service (Claessens et al., 2003). This is especially relevant for payment systems, where benefits and thus demand increases as the pool of users expands. High fixed costs can trap a small financial system at a low level equilibrium because of the system's inability to reap the necessary scale economies and network externalities.

In sum, fixed costs can constitute an important limitation to outreach in the provision of payment and savings services and, hence, a key barrier to the broadening of access to these services. Profitable and sustainable financial intermediaries have to exploit scale economies either through sufficiently high-volume *or* high-value transactions, but not necessarily through both. In particular, the competitive environment could be such that deposit-taking institutions could stay in business without much outreach efforts, by specializing in large-value payments and savings services.

Because of scale economies and network externalities, problems of access to payments and savings services in many developing countries are related to the oft-found triple problem of smallness—small transactions, small financial institutions, and small market size. Unless a way is found to raise transaction volumes to seize scale economies, low-income clients with the need for small and few payment and savings transactions would not constitute a profitable clientele for financial service providers. A small, community-based financial institution might not be able to spread the fixed costs of its physical installations, technological platform, legal and accounting infrastructure, etc. over a sufficiently large business volume to be profitable. Similarly, a bank might not be willing to open a branch in smaller towns if there is not sufficient market potential to cover its fixed costs of setting up shop. Finally, a small market might be able to support only a few viable financial institutions (i.e., institutions that reach an efficient scale), thereby having to forego the type of competition that—as we argue below—appears crucial to foster the broadening of access.

It is useful to think of the process of supply-originated broadening of access to payments and savings services as driven either by changes in "state variables" or by idiosyncratic cost

management for a given level of state variables.[6] While idiosyncratic cost management refers to actions within the realm of individual financial institutions, we treat as state variables those that are largely outside the control of the managers of financial intermediaries and that change slowly over relatively long time periods, including the following: market size, macroeconomic fundamentals, available technology, the average level and distribution of per capita income, and system-wide costs of doing business related, for instance, to the quality of transport and communication infrastructure, the effectiveness of the contractual and informational frameworks, and the degree of general insecurity associated with crime, violence, terrorism etc. We argue that the intensity of competition can vary for a given level of state variables, but that competition is also a key driver behind changes in state variables over time. An oligopolistic market structure, for example, can result in dominant banks successfully opposing any improvement in the contractual and informational frameworks as this would reduce their rents.

The interplay between state variables and idiosyncratic cost management in determining supply outreach for payments and savings services is illustrated in Figure 1. It shows, via iso-profit curves, the combinations of transactions value and transactions number of payment and savings services that yield the same profit for a given financial institution. The horizontal axis measures the number of transactions which we assume, for ease of presentation, to increase only with the number of clients—i.e., we hold the number of transactions per client constant. The vertical access measures the "typical" value (or the mode) of the payment or savings transaction handled by the institution.[7] The iso-profit curves are drawn for a given level of state variables and are downward sloping because an increase in the number of transactions (clients) is needed

[6] It is of course somewhat arbitrary to define the set of state variables, not least because the definition is not independent of the relevant time horizon.

[7] This is equivalent to assuming that the vertical axis measures the average value of the transactions of a client and that such value increases only with the value of each individual transaction, so that for ease of argument the number of transactions per client is held constant.

for a financial entity to make the same profit while moving towards a lower typical transaction value. The curvature of the iso-profit curves depends on the characteristics of the production function for payments and savings services. The iso-profit curves are parallel, with curves more distant to the origin representing higher profits. A movement *along* an iso-profit curve towards a larger number of clients represents, by definition, a supply-induced broadening of access.

It might be not that easy, however, for financial institutions to trade-off smoothly (beyond a certain range) a higher transaction value for a higher number of clients while still making the same profit. Discontinuities along this trade-off could arise if the supply of large-value payments and savings services is not generated by the same production function as the supply of low-value/high-volume services, and changes from one production function to another imply switching costs. Technology, marketing and costumer service approach, location, size, configuration of bank offices, and even the profile of staff are all a function of the type of service produced and switching to a different technology, marketing, etc. to reach a new clientele is costly. The case of two production functions is depicted in Figure 1, where financial institutions cannot move from point A to point B without incurring switching costs, even if both yield in the end the same profit. This, in turn, might be a barrier to move down-market to expand access.

Discontinuities in the iso-profit line explain not only why financial institutions cannot move freely along a given curve beyond a certain range of values. They also can also help explain why financial systems in many emerging economies show a clustering of financial institutions at the North-West corner of the space in Figure 1, with few if any institutions at the South-East corner. That type of clustering reflects a situation where banks cater mainly to large value clients and are unmotivated to incur the costs to change their production function to move

down-market to a point like B. It also reflects a situation where newcomer banks do not enter, for one reason or another, business niches located in the South-East corner of Figure 1.

Why would that clustering equilibrium arise even if the production technology is available to make point B feasible and equally profitable as point A (as is suggested in Figure 1)? One obvious explanation would be that profits in that type of banking system are not under strong competitive pressures. Due to insufficient market contestability, payment and savings service providers could make money comfortably by targeting larger firms and wealthier households and would have little incentive to manage their costs efficiently to reach out to new clients by, say, re-engineering their internal processes and services and optimizing on the use of available information technology. Insufficient competitive pressures may also be associated with barriers that deter newcomer service providers—who by definition do not incur switching costs—from entering the business of payments and savings service provision to target the smaller firms and poorer households.[8]

The clustering of payments and savings service provision in the North-West corner of Figure 1 could also occur—even if competition was strong—because of low levels in certain state variables. For instance, a small market size and/or lack of technology to reduce costs sufficiently could hinder supply outreach. In this case, the iso-profit curves would not extend into the South East corner of the graph.

Even in the absence of discontinuities, the supply-driven broadening of access to payments and savings services could conceivably encounter difficulties if iso-profit curves were "too flat." Relatively flat curves would mean that to remain at least equally profitable a financial institution would have to increase the number of transactions (clients) very substantially for each

[8] Given the fixed cost character of the switching costs, bigger banks are in a better position to incur these costs than small banks.

small decrease in the typical value of the transactions it handles.[9] If iso-profit curves were flat, increased market contestability would be insufficient to promote supply-induced broadening of access, as that flatness would be a reflection of state variables. A significant growth in market size, a technological breakthrough in information and communications technology,[10] a noticeable improvement in road or telecommunication infrastructure, or a palpable reduction in general insecurity would be required to create greater curvature in iso-profit lines, thereby facilitating supply-induced broadening of access.

While Figure 1 illustrates different supply constraints to access, we have to take into account demand side constraints. Further, the analysis up to now has focused on individual institutions, while ultimately we are only interested in institutions that work on the margin, i.e. with the most marginal customers. Therefore, we now turn to a discussion of demand-side constraints to outreach before combing aggregate demand and supply side analysis to derive the access possibilities frontier.

### *(b) Economic and non-economic demand factors*

Price and income level are the salient economic determinants of the demand for payments and savings services.[11] Economic development and the associated rise in per capita income increases the need for more and more sophisticated versions of these services. However, demand is not only driven by economic but also by socio-cultural factors. Further, we have to isolate

[9] Whether flat iso-profit curves are not just a theoretical possibility but also a phenomenon of empirical relevance is a question beyond the scope of this paper. To our knowledge there is no study that investigates this empirical issue.

[10] The IT revolution has allowed new and cheaper forms of delivering and processing financial services, including via the phone and internet. See Claessens, Glaessner, and Klingebiel (2002) for a discussion of the potential role of e-finance in emerging markets.

[11] Peachey and Roe (2004) show that the increase in the share of population with bank accounts across OECD countries over the past 50 years can be attributed to the entry of women into the work force and the associated higher disposable household incomes.

pure demand factors from demand reductions that are due to the expectation of supply constraints.[12]

In the following, we will distinguish between two demand curves—a potential demand curve, driven purely by economic factors, and an actual demand curve, that might be below the potential one due to non-economic factors. We can write the potential (individual or aggregate) demand as D* = f{income, price}, with demand increasing in the first argument and decreasing in the second. The actual demand can be lower than potential demand for a given price and income level, due to self-exclusion arising from such non-economic reasons as financial illiteracy and ethnic or religious factors.[13] Taking into account these non-economic demand factors, we can write actual demand (individual or aggregate) as: D= f{income, price, financial illiteracy, cultural barriers}, with demand decreasing in the third and fourth arguments. Abstracting from the theoretical possibility that illiteracy of cultural biases might lead to over-demand of financial services, actual demand will be bounded by potential demand.

### *(c) The Access Possibilities Frontier for payment and saving services*

Aggregate supply and demand determine the bankable population, i.e. the share of population a financial system can reach with payment and savings services. We define the Access Possibilities Frontier for payment and saving services as the maximum share of population that could be served by financial institutions, for a given set of state variables. Figure 2 combines the demand and supply analysis of the previous two sections into the traditional price-quantity graph. It assumes away aggregation problems; specifically while there may be

[12] This would include lack of demand due to expected discrimination-based rejection.

[13] Financial illiteracy might lead to lower demand for services than predicted by economic capacity to pay as people might not know of the existence or affordability of specific payments or saving services, partly due to a generally low level of education.

different production functions in the supply of high-value services versus low-value/high-volume services at the level of individual financial institutions, Figure 2 assumes that we can speak cogently of only one type of payments (or savings) service at the aggregate level, i.e., that the mentioned differences in production do not prevent a smooth blending of individual supply curves as they are aggregated into the overall supply schedule.[14]

In line with the assumptions discussed at the beginning of this section, the price measured along the vertical axis is a fee, defined on a per-transaction basis. We assume further that the fee is flat, i.e., independent of the size of the transaction. On the horizontal axis, rather than the quantity of service transactions, we plot the share of the population (households and firms) engaging in payments and savings transactions. We order the population share along the axis starting with agents that engage in transactions that are large in value and number and moving towards agents with transactions of increasingly lower value and number. To simplify this ordering, we make two further assumptions. First, that clients with a high transactions volume are also characterized by transactions of high value. Second, that while the volume of transactions per period and value of each transaction vary across agents, each agent engages in a fixed number of transactions of equal value per period, independently of the fee.[15] These assumptions imply that movements of the equilibrium outcome towards the right of the graph unequivocally represent a broadening of access to payments and savings services, as marginal customers that demand low-value/low-volume transactions enter the financial system, raising the share of the banked population.

[14] Alternatively, we could have assumed discontinuities, which could result in multiple equilibria. For the ease of discussion, we exclude this possibility in our analysis.

[15] This implies that when the fee falls, individuals do not react by demanding more transactions, but rather more individuals will demand. Or, in the terminology of the trade literature, changes in the fees have effects only on the extensive not on the intensive margin.

Let us first discuss the potential and actual demand curves for payments and savings services as function of its price and for a given set of state variables (which, as noted earlier, are beyond the influence of individual service providers and include market size, macroeconomic fundamentals, the average level and distribution of per capita income, available technology, physical infrastructure, the contractual environment, and the degree of general insecurity). The demand reflects willingness to pay and we expect customers with demand for large-value/high-volume transactions to have a higher marginal willingness to pay than customers with few and small transactions. A higher price thus reduces the share of the population that will demand payment and savings services, thus resulting in a downward sloping demand curve. As the price falls, marginal users (i.e., smaller firms and poorer households) add their demands for typically lower-value transactions. Curve D* denotes the potential demand as function of economic factors only, while D denotes the actual demand, as predicted by both economic and non-economic factors, including self-exclusion. The distance between the two curves denotes self-exclusion and is not necessarily constant across the population spectrum. However, if one conjectures that self-exclusion and illiteracy are inversely correlated with income, the gap between potential and actual demand opens up as marginal households and firms are added to the demand, as depicted in Figure 2.

We now turn to supply. As the fee increases, it becomes profitable for service suppliers to reach a larger share of the population, by targeting agents (households and firms) characterized by increasingly lower-value transactions. This results in an upward sloping supply curve. The higher steepness of the supply curve in the outer regions could be explained by either switching costs or flat iso-profit curves, as discussed above. The curve S* denotes the potential supply, representing the share of population that can be serviced for a given price by efficient

financial service providers, given the state variables. Note that the S* curve assumes efficiency in supply, i.e. that service providers maximize profits *for a given price* and are, thus, located on the highest feasible iso-profit curve. The potential supply, S*, differs from the actually observed supply curve, S, which denotes the share of population that currently existing financial service providers are actually willing to cater to for a given price. As illustrated in Figure 2, S is to the left of S*, implying that the former is a higher cost supply due to some form of inefficiency, including distortionary regulatory policies, an excessive number of banks for the size of the market (leading to a failure to achieve economies of scale), or lack of adequate market contestability.

We are now ready to analyze various key points in Figure 2. Point I, the intersection of potential demand D* and potential supply S* schedules, denotes the access possibilities frontier, i.e. the maximum outreach point for payments and savings services that can be reached in a country's financial system, given the state variables. Point I is, thus, the constrained optimum, associated with a bankable population at point A on the horizontal axis. Hence, 1-A is the non-bankable population, that is, the share of population that cannot be provided with market-based payments and savings services, given the state variables.

The access possibilities frontier can be moved by shifts (or changes in curvature) in the potential supply and the potential demand curve as a result of changes in state variables. An outward shift of the potential demand curve due, for instance, to a higher level of per capita income increases access, as measured by the equilibrium share of bankable population. Similarly, a downward shift in the potential supply curve due to improvements in a country's infrastructure or institutional environment will result in an expansion of the bankable population.

Starting from the Access Possibilities Frontier and the bankable population (Points I and A), we can identify three different types of access problems. First is the access problem implied in point II, which denotes a lower than potential possibilities frontier as a result of non-economic factors that lead to self-exclusion of agents in the demand schedule. In this case, the equilibrium level of banked population (B) is lower than the bankable potential, given state variables. The A - B distance can be interpreted as a measure of a demand-driven access problem.

A second type of access problem is illustrated by points III or IV. Both are points of constrained sub-optimality, i.e. an inefficient (high cost) supply leads to an equilibrium where the banked population is lower than the bankable population, given the state variables. The access problem is worse at point IV which, in addition to supply inefficiency, reflects also self-exclusion. The distances A – C and B – D are measures of access problems driven purely by supply inefficiency, as they hold the type of demand (potential or actual, respectively) constant. The distance A – D is a measure of an access problem caused by a combination of demand self-exclusion (A- B) and supply inefficiency (B – D).

The third type of access problem would obtain if the bankable population associated with point I is "too low" relative to countries with comparable levels of economic development. This situation could arise, for example, if the country in question lags behind its comparators in certain state variables (say, higher level of general insecurity or weaker informational and contractual environments).

***d) Policies to Foster Access to Savings and Payments Services***

In order to design effective policies to expand access to payments and savings services, policymakers have to be aware of which of the three access problems is the most binding one

and/or can be addressed in the most effective way. More specifically, policy makers should try to understand if the access problem is mainly because their financial system is inside the possibilities frontier, and why, or mainly because the achievable frontier is too low compared to similar countries and why. The latter problem will generally call for *market-developing* policies that expand the possibilities of outreach through structural reforms that improve institutions and other state variables, thus moving the frontier outwards over time. The former problem rather calls for *market-enabling policies* that, depending on the underlying reason, would put emphasis on raising financial literacy to mitigate self-exclusion or providing incentives to financial service providers to operate more efficiently. [16]

Consider first *market-enabling policies*. While the IT and communication revolution allows financial institutions to improve cost management, the motivation to seize such opportunities typically comes from the pressures of competition. In effect, competitive pressure and the search for profits are key factors behind such examples of outreach as offering of services that are tailor-made for low-income clients (e.g., simple debit accounts at lower costs than regular checking accounts) or the use of mobile branches or cell phone banking to reach populations in remote areas at low costs. Policy makers have an important role to play in providing financial institutions with the necessary incentives to expand outreach through different regulatory policies and actions.

If the lack of innovation and the corresponding increase in outreach is due to imperfect competition, policy responses should aim at enhancing market contestability, so that service providers—existing or new—are under competitive pressures to achieve cost minimization through efficiency gains, holding the state variables constant. Allowing foreign bank entry or

[16] Compare a similar classification in Porteous (2004).

avoiding overly high licensing and minimum initial capital requirements can foster contestability and more specifically the emergence of financial institutions catering to the needs of low-value customers.[17] Preventing the establishment of closed clubs (such as payment networks limited to large banks) can keep a system contestable and thus competitive.[18] Looking beyond the commercial banking system and allowing competition from the non-bank financial sector can be important. For instance, preventing payments infrastructures from becoming inaccessible to financial institutions outside a restricted "club" and ensuring that retail payments infrastructures are open and interoperable can help the financial system cater to marginal customers at low cost. If inefficiency reflects instead the inability to capture economies of scale because, say, the number of banks is too large relative to the size of the market, policy responses should first aim at the consolidation of the banking system, through the exit of unviable institutions via orderly closures and mergers.

Regulatory policies can have a profound impact on financial institutions' costs and efforts to increase outreach, even if they do not target directly the competitiveness of a financial system. On the negative side, high compliance costs with unduly complicated "Know Your Customer" (KYC) and anti-money laundering (AML) regulations may prevent financial institutions from moving to the frontier (or prevent the establishment of institutions at the frontier) to service marginal customers (Claessens, 2006). On the positive side, relaxing branching restrictions by allowing financial institutions to offer limited services through non-financial correspondents can significantly reduce the fixed cost element of financial service provision and thus problems of

[17] The deregulation wave in OECD over the past 30 years and the consequent increased competition led many financial service providers to look for new markets, effectively expanding the banked population (Peachey and Roe, 2004).

[18] For cross-country evidence on the link between regulatory policies and firms' access to finance, see Beck, Demirguc-Kunt and Maksimovic (2004). For cross-country evidence on the link between market structure, regulatory policies and the competitiveness of a financial system, see Claessens and Laeven (2004).

diseconomies of scale in remote and small market places.[19] A further step would be to allow the use of public post office networks as platform for service provision by different financial institutions, as in India and South Africa (see World Bank, 2004 a, b). This cannot only help overcome problems of scale economies, but address concerns of competition compared to the situation where only one institution is allowed to use post offices as correspondents as in Brazil or a situation where one financial institution dominates service provision in remote areas, as is the case in large parts of Sub-Saharan Africa.

Sometimes, the role of government in fostering access might have to go beyond competition policies and take the form of "affirmative regulatory" policy. Examples include the moral suasion exercised by authorities to make South African banks introduce the Mzansi (basic transaction) account or make British banks introduce the Basic Bank Account (BBA).[20] Inducing banks to share or ensure interoperability of payments infrastructures (including ATM networks) can help avoid undesirable competition on access to infrastructure while enhancing desirable competition on price and quality of service, thereby facilitating the achievement of cost-reducing scale economies and lowering entry barriers to new financial service providers.

Market enabling policies can also be important on the demand side. Financial literacy programs or financial products tailored to specific groups can increase access to and use of financial services by bringing actual demand closer to potential demand and move a financial system closer to the frontier.[21] The government, in its role as financial market participant, can also have an important role. For instance, in 1999, the U.S. Treasury started a program to make

[19] In Brazil, for instance, some of the largest banks have expanded their network through correspondent agreements with the Post Office, lottery shops and supermarkets (World Bank, 2005).

[20] See Napier (2005) for a discussion of the Financial Access Charter in South Africa. This is very different from Universal Service Obligations, as legislated in several European countries (Kempson, Atkinson and Pilley, 2004), which require banks to open a bank account for anyone. Such obligations can force banks to move access beyond the constrained optimum and raises questions on sustainability and efficiency.

[21] Compare Financial Services Authority (2000) and the literature quoted therein.

all federal benefit payments though electronic transfer accounts, using subsidies to encourage banks and beneficiaries to open such accounts (Claessens, 2006).

Market-enabling policies, however, find their limits in the state variables. Deficiencies in state variables require more profound structural reforms, mostly outside the reach of even financial sector policy makers. We denote such reforms as *market-developing* policies. One often neglected area are distortions in the input markets for financial service provision, including labor and communication markets. A deficient phone infrastructure and lack of competition in the telephone market can drive up costs for financial institutions.[22] A deficient transportation infrastructure can increase the cost of outreach into more remote areas, even through innovative techniques such as mobile banking. Generally high costs of doing business, due to widespread corruption, insecurity, or even deficient electricity provision, can drive up the fixed cost element of financial service provision with negative repercussions for outreach to low-value/high-volume customers. Finally, economic size can be a restricting factor as discussed above. Small economies should therefore put a premium on encouraging entry of foreign banks that are able to reap benefits of scale economies across subsidiaries in different countries, on integrating their local markets with international financial markets, and on allowing their citizens access to financial services across borders.[23]

[22] Compare the detailed discussion for the case of South Africa in World Bank (2004b).

[23] For a more detailed discussion on policy options for small financial systems, see Bossone, Honohan and Long (2002). Ultimately, the last recommendation leaves open the question to which extent domestic financial systems are viable at all in small developing economies and agents should not rather rely on providers abroad for certain services. Even in the extreme, some local markets will be needed, especially for local-currency denominated financial services, which are important not just for financial access but also to avoid financial vulnerabilities stemming from currency mismatches.

## III. The Analytics of Access to Credit

This section expands the concept of the Access Possibilities Frontier to lending services and introduces risk as the key feature that distinguishes credit services from savings and payment services. While we will be referring to lending services in general, most of our analysis focuses on a given loan product rather than on aggregate lending services, as different types of loans entail different technologies and, hence, different production functions, implying switching costs in moving from the provision of one type of lending to another.

### *(a) Risk as barrier to outreach in credit services supply*

In addition to costs, the outreach in the supply of credit services is constrained by risks, especially default risk. These can be either borrower-specific or systemic. For the purposes of this paper, we define as systemically originated credit risk that which is non-diversifiable within a given domestic jurisdiction and as a consequence affects all credit contracts therein.[24]

Systemic risk typically stems from high macroeconomic uncertainty (e.g., significant volatility in the rate of inflation, the terms of trade, the real interest, and the real exchange rate), weaknesses in the contractual and informational environment (e.g., poorly defined and difficult to enforce creditor rights, deficient accounting and disclosure practices, and lack of a well functioning credit bureau), or geographical limitations (e.g., a small country prone to flooding or hurricanes).[25] Regardless of its origin, systemic risk hinders the supply of credit because it raises the default probability and the loss given default for all credit contracts written in a given jurisdiction. This leads to a higher than otherwise cost of funds and, hence, a higher floor for the

[24] In our definition, therefore, systemic risk might be diversifiable through international financial markets.

[25] De la Torre and Schmukler (2005) analyze the role of systemic risk in credit contracts and show that the terms of loan contracts—particularly duration, currency of denomination, and jurisdiction—are adapted so as to cope with systemic risk.

interest rate required by a creditor to make a loan. As systemic risk increases, it enlarges the set of borrowers/projects that find the cost of credit unaffordable and are thus priced out of access to credit.

Idiosyncratic credit risks are specific to individual borrowers/projects and are, hence, not correlated with systemic risk. As a result, the cost of finance and/or availability of credit services will differ across debtors/projects depending on their differences in idiosyncratic riskiness, and is priced as a spread over the interest rate floor set by systemic risk. Importantly, however, the ability of the lender to manage idiosyncratic risk is influenced by the systemic risk environment as we will discuss below. Two factors are particularly important in explaining the differences in interest spreads across debtors (for a given type of loan) that are induced by idiosyncratic risk: agency problems and limits to the diversification of risks that are not related to agency problems. Let us explain these factors in some detail.

Consider first the constraints on the ability to reduce non-agency related risk (for a given return) through diversification. Idiosyncratic risk would in principle be diversifiable or insurable (and thus not priced in the interest rate spread) in an idealized world where markets are complete. The limits to idiosyncratic risk diversification observed in the real world are thus a reflection of some form of market incompleteness, including the lack of sufficient markets for hedges and other insurance products. If unable to diversify non-agency related risks in a competitive market, risk adverse creditors will include a risk premium in the lending interest rate. Such risk premium increases the lending interest rate beyond the level necessary to cover the creditor's marginal cost of funds plus the cost she has to incur in order to provide the credit service. In all, high risk premiums undermine credit supply outreach to the extent that they render lending interest rates unaffordable for certain borrowers. The lack of agricultural lending

in many developing countries has often been explained by the inability of financial institutions to diversify the high risk stemming from agricultural activity and therefore agricultural lending.

Consider now agency problems. They give rise to a misalignment of debtor incentives vis-à-vis the interests of the creditor. The most studied source of agency problems is information asymmetry, whereby the debtor is privy to relevant information about herself and her project that the creditor may not be able to secure or only at a prohibitively high cost, and which can lead to two conceptually distinct sources of credit risk: adverse selection and moral hazard.[26] Although higher agency-related idiosyncratic credit risk can be compensated by increasing the risk premium as discussed above, such an increase can lead to adverse selection and exacerbate the moral hazard problem.

Regarding adverse selection, a higher risk premium would tend to attract riskier borrowers, i.e., borrowers more likely to default in bad states of the world, to the pool of loan applicants while the creditor may lack sufficient information to sort them ex-ante. As a result, instead of raising the risk premium, creditors try to mitigate adverse selection through non-price screening devices (such as collateral or character assessment) or simply deny credit to loan applicants that are not favorably screened through such devices. Moral hazard, for its part, arises from the borrower's incentive to use the proceeds of the loan in endeavors that are riskier than those specified in the credit contract, while being able to conceal such behavior from the creditor. By taking on greater risks, the debtor can capture the windfall gains in the good states of the world while limiting her loss to her investment in the project (assuming limited liability) in the bad states of the world. An increase in the non-agency related risk premium, furthermore, can intensify moral hazard by further exacerbating the debtor's incentives to divert the loan

[26] The classic article on the effects of information asymmetry on credit supply is Stiglitz and Weiss (1981). Though conceptually distinct, adverse selection and moral hazard are difficult to distinguish empirically, as discussed by Karlan and Zinman (2004).

resources to riskier uses. Thus, rather than increasing the interest rate, creditors may again resort to non-price devices to mitigate moral hazard or simply curtail the quantity of credit supplied.

Figure 3 allows us to illustrate more formally the way in which systemic and idiosyncratic risks influence the cost and availability of credit supply. The figure is drawn for a given type of loan (say a consumer, mortgage, working capital, or investment loan) supplied by financial institutions in a given jurisdiction. The lending interest rate, i, is measured along the horizontal axis, and the expected return to the lender, r, is measured along the vertical axis. In an ideal world devoid of transactions costs and risks, i and r would coincide and their relation would take the form of a straight 45 degree line from the origin. In the real world of market frictions and risks, however, there is a wedge between the interest rate charged by the creditor and his expected return. We will now show how the three different risk categories lead to a deviation of the lender's return from the lending interest rate and potentially to credit rationing.

Take first country-level systemic risk, illustrated by a parallel line to the 45 degree line, and with this difference denoting the country risk premium, c. Denoting with $i_{mc}$ the marginal cost of funds for the creditor bank—which, due to arbitrage, is assumed to be equal to the interest rate paid on short-term government debt securities—we can establish a floor for lending interest rates and a first wedge as $r_{mc} = i_{mc} - c$. The country risk premium, c, can be explained by the different elements of systemic risk explained above—macroeconomic uncertainty, deficiencies in the contractual and informational frameworks and geographic limitations. This systemic risk is common to all credit contracts in a country and cannot be diversified away within that jurisdiction.

Take next the non-linearities due to scale and agency problems, illustrated by curve I.[27] The nonlinear wedge between the 45 degree line and curve not only implies that the default probability increases with the lending interest rate, causing r to rise less than i. It also implies that, as the lending rate increases beyond a given threshold, denoted in Figure 3 by i*, the expected return begins to decrease. Thus, at (i*, r*), the marginal revenue to the creditor due to a contractual increase in the lending interest rate is fully offset by the marginal expected loss due to a higher probability of default. This nonlinear relationship can be explained by agency problems, for the reasons discussed above. But it could also be rationalized without resort to agency problems. For example, as shown by Williamson (1987), a similar nonlinear wedge can arise from the existence of transaction costs in lending combined with uncertainty in debtors' revenue streams.[28]

Consider finally the incapacity to diversify non-agency related risk. Curve I is drawn after subtracting from the interest rate any non-agency related idiosyncratic risk premium. Hence, along curve I, a higher interest rate incorporates only higher lending costs due to smaller scale and/or higher agency-related idiosyncratic risk. Curve II, on the other hand, takes into account the risk premium and, hence, is always to the right of curve I, with the vertical distance between the two curves measuring the premium charged by creditors for non-diversifiable risk. To the extent that the risk premium increases with the level of the lending rate (reflecting the increase in the ex-ante probability of default), curve II would be flatter than curve I and would have a lower flexion point, as drawn in Figure 3. Note that the widening of the wedge between i and r as i increases is common to both curves. This is because the probability of default rises

[27] Note that this curve starts at $i_{mc}$, as no loan would be made by a creditor at an interest rate below his marginal cost of funds or opportunity cost of lending.

[28] In such a situation, an increase in average transaction costs increases the loan interest rate the lender has to charge in order to recover her costs and thus increases the probability that the borrower may not be able to repay due to a negative output shock, independently of her willingness to repay.

with the lending interest rate, independently of the reasons (costs, risk-adjusted profits, or risk premium) that push that rate up.[29]

Figure 3 allows us to focus on three types of distinct but interrelated state variables that constrain the supply of credit services. The first state variable is macroeconomic risk which, as noted, is assumed to be fully summarized in the interest rate paid by the sovereign on its debt. Higher macroeconomic risk raises the opportunity cost of lending, i.e., $i_{mc}$, for all creditors in a given jurisdiction, thereby increasing the floor for the minimum interest rate that can feasibly be charged on loans to the private sector. The higher this state variable, the larger the set of households and firms that will be priced out of the credit market.

The second state variable is the quality of the contractual and informational environment. The lower such quality, the more difficult it will be for creditors to mitigate information asymmetry and contract enforcement problems in "arms-length" credit transactions and, hence, the more they will resort to non-price screening and monitoring mechanisms, such as highly personalized "relationship lending" and/or heavy reliance on fixed (mainly real estate) collateral from debtors. The lower the quality of the contractual and informational environment the flatter Curves I and II in Figure 3 and the wider the wedge between i and r that cannot be reduced through idiosyncratic effort, i.e. effort by the individual lender. As a result, the lower such quality, the larger the set of debtors that, despite having viable projects, i.e., projects that debtors

[29] The supply function for loans to the private sector can be easily derived from Curve II in Figure 3, by noting that loan supply, $L_S$, is a positive function of expected return, r, which is in turn a function of the lending interest rate. Formally, $L_S = f\{r(i,\ldots)\} = g\{i,\ldots\}$, where $\delta g/\delta i > 0$ up to a certain threshold. Moreover, the non-linear relationship between nominal interest rate and expected return to lender can result in backward-bending supply curve - $\delta g/\delta i < 0$ - and credit rationing, as shown by Stiglitz and Weiss (1981).

would be willing to finance out of their own resources if available, will have no access to external finance.[30]

The third state variable is given by systemic limits to the diversification of non-agency related risks. As noted, such limits reflect undiversified productive sectors and financial underdevelopment, manifested in missing markets for certain loan products as well as for hedges and insurance products.[31] The narrower the scope to diversify away idiosyncratic risk, the higher the risk premium that will be incorporated in the interest rate for any given type of loan. High risk premiums further enlarge the set of potential debtors that are priced out of the credit market. High risk premiums, moreover, exacerbate agency problems, further enlarging the set of debtors that is rationed out of the loan market.

As state variables are out of the control of individual lenders, their main job consists of managing idiosyncratic risks and costs so as to make a profit. Lenders must develop loan technologies that enable them to choose such debtors that, for any given lending interest rate, will give the highest risk-adjusted expected net return. To select such debtors for a given type of loan, the creditor must compare: (i) the all-in costs of lending to different debtors/projects; (ii) the differences in expected returns (capacity to pay) across such debtors/projects that the lender considers to be equally risky; and (iii) the differences in risks and willingness to pay across such debtors/projects that the lender considers to be of equal expected returns (or capacity to pay). These comparisons require approaches that differ across different types of loans, ranging from

[30] De la Torre and Schmukler (2006a) use this "agency wedge" as their working definition of an "access to finance problem." They note that, according to this definition, an access problem is reduced to the extent that the *share* of viable projects that are able to obtain external finance increases.

[31] The missing market problem, moreover, can be exacerbated by deficiencies in the other two state variables. For example, high macroeconomic volatility can itself be a root cause of missing derivatives markets or missing markets for long-duration local-currency debt. The problem of missing markets not only limits the ability of creditors to diversify risks but also makes it more likely for debtors to incur currency and duration mismatches in their balance sheets, which, in equilibrium, can be an optimal outcome in high systemic risk environment (De la Torre and Schmukler, 2005).

scoring methods based on the law of large numbers for the more homogeneous loan products (such as credit card loans, micro loans, or even mortgage loans), to the low-cost low-risk lending to the most reputable, creditworthy, and typically large debtor firms, and to the more intense and personalized approaches in lending to the opaque and clearly more heterogeneous SMEs. After having selected the debtor/project that will be offered a loan, the lender has to monitor the debtor's behavior and has to decide what to do in case of default, i.e. to write-off the loan or engage in post-default value recovery efforts. Finally, the scope for optimization that the lender will have in managing lending costs and risks will be constrained by the state variables. The weaker the state variables, the less the maneuvering room for credit supply optimization. Given the constraints, a prudent lender will rather not offer loans at a higher interest rate if she is not reasonably sure of her ability to appropriately measure costs, sort out risks for a given expected return, and identify expected returns for a given risk.

There is a correlation, but far from perfect, between cost and risk management. Risk management activities aim at better measuring and controlling default risk exposure in the production of credit services, while cost management activities aim at minimizing costs for a given unit of output. A given management activity may accomplish both aims. For instance, the introduction of computer-based statistically-driven scoring models to measure and control default risk in consumer or micro credit may also be seen as a cost management activity to the extent that it enables a reduction in the number of loan officers and branches (given that scoring methods could be applied via e-banking). By contrast, some activities may only affect directly cost management (e.g., upgrading the IT hardware) or only risk management (e.g., defining the types of assets that would be accepted as collateral).

Lenders must confront complex interactions between costs, returns, and risks. Clients that are equally risky may be associated with different expected returns (or capacity to pay). Clients with similar payment capacity may have different risks and willingness to pay. Clients with equal risk and capacity to pay may represent different costs for the bank. Costlier clients are often, but not always, also the riskier and smaller clients. For instance, small-scale enterprises demand smaller loans, are less transparent, and can often offer fewer risk-alleviating remedies such as fixed collateral and guarantees. In contrast, the marginal borrower in long-term investment loans is not necessarily small, as the inability of lenders to offer longer-term investment loans is mostly due to a missing market problem arising from systemic risk considerations and independent of loan size.

Be it as it may, holding the state variables constant, a supply-driven broadening of access to credit services can be generally defined as the incorporation of additional, marginal borrowers into the credit circuit for any given lending interest rate and loan type. The marginal borrower would generally be the one that, in the view of the lender, will give a risk-adjusted return that is just sufficient to cover costs. Compared to our previous analysis of savings and payment services, however, where the marginal client was identified as the smaller firm and poorer household (characterized by low value transactions), the marginal client in the supply of credit services is not necessarily defined along one dimension. It may be the costlier client to serve (due to opacity, riskiness, or smallness); it may be the client with a lower expected return among equally risky debtors; or it may be the client with the higher risk among debtors with equal expected returns.

We can now take the analysis one step further by considering the supply and demand for credit services, which allows us to develop the Access Possibilities Frontier for loan services in a similar fashion as was done in Section II for payment and savings services.

***(b) The Access Possibilities Frontier for credit services***

As we combine demand and supply for credit services, our focus on access means that we are interested in the marginal borrower/project. However, as noted, identifying the marginal borrower/project is not a simple matter, given the complex interplay of costs, returns, and risks in credit markets. Because the multi-dimensionality of lending constraints is captured in one price—the interest rate—there is not a unique combination of cost, size, risk, and return of projects that maps one-to-one to interest rates. Furthermore, given information asymmetries, the ranking of borrowers according to, say, expected returns or riskiness, will be different depending on whether it is done by the lenders or the debtors. Thus, demand and supply cannot be drawn in the same space that has the interest rate in the vertical axis and a given ordering of borrowers in the horizontal. This is different from payment and savings services, where transaction costs and size of transaction can be mapped directly into the price and thus into demand and supply. Further, changes in the price, i.e. the interest rate, cause changes in borrowers' behavior beyond ability to pay (unlike in the case of payments and savings services) as discussed above. Due to asymmetric information, however, these changes are not obvious to the lender; while she might know the overall riskiness of the portfolio, she does not know the riskiness of individual borrowers and their precise change in behavior as the interest rate changes. In all, the multidimensionality of lending constraints and the reaction of borrowers to changes in interest rates make it impossible to combine, in the traditional way, demand and supply into one graph.

We will therefore take an alternative, two-step approach, first defining demand for a given loan product and given a certain $i_{mc}$. Taking this defined universe of potential loan applicants, we will then illustrate supply as a function of the marginal interest rate, introduce demand in an indirect and non-traditional way, and derive an Access Possibilities Frontier that will help us determine the share of borrowers in a pool of loan applicants that are potentially bankable and actually banked, as function of the interest rate.

Consider first demand, illustrated in Figure 4. The vertical axis denotes the nominal marginal interest rate charged by lenders, while the horizontal axis has the number of loan applicants, who are ordered according to the inverse of the expected return of the project that they advertise to lenders.[32] Hence, as we move along the horizontal axis and to the right we encounter applicants with lower and lower expected returns. For simplicity, we assume that each applicant demands only one loan, although we allow loan sizes to vary across applicants. We can thus be sure that the downward slope in the demand curve implies that a decrease in the lending interest rate adds a new applicant with a marginally lower expected return to the pool of demanders (rather than adding a new loan by a borrower already in the demand pool).[33] We draw three different demand curves in Figure 4, corresponding to (i) demand in a world without market frictions, defined as a world where borrowed resources are treated by the debtor as if they were his own resources; (ii) demand taking into account self-exclusion due to cultural barriers and financial illiteracy; and (iii) demand taking into account limited liability, asymmetric information, and the resulting agency problems. All three demand curves are contingent on the

[32] To simplify things, we assume that the expected return that borrowers advertise to lenders is equal to the one they expect to get. As will become clear later, however, this advertised expected return may be higher than the return that borrowers expect the project itself to deliver.

[33] For ease of graphic illustration, we are using a straight line for the demand curve. In reality, this does not have to be the case.

investment opportunities available in the economy—the lesser the investment opportunities the more the demand curve would be to the left of the graph.

Take first demand corresponding to a world without market frictions, represented by Curve I, which serves as benchmark for the further analysis. This illustrates demand for loans that borrowers would treat like their own resources or, alternatively, demand that involves only those projects that the debtor would undertake with her own resources if she had them. Imposing (for ease of argument) a zero-profit condition on borrowers, this would imply that the marginal borrower with expected return x is willing to pay an interest rate of up to x% on her loan. Demand is cut off at $i_{mc}$, i.e., no borrower will demand loans for a project with expected return lower than $i_{mc}$, the return that could be earned by investing in debt securities of the local sovereign.

Not all potential borrowers with profitable investment projects will demand loans, however. Cultural barriers to incurring debt or financial illiteracy resulting in ignorance about borrowing possibilities might result in self-exclusion and, thus, a loan demand as represented by Curve II, to the left of Curve I.[34]

While cultural barriers and financial illiteracy might reduce demand, incentives arising from limited liability cum asymmetric information are demand-increasing—they result in a demand for loan resources that exceeds what the borrower would invest in the project if she owned such resources. Limited liability leads to an asymmetric participation of the borrower in the returns of a project—the gains from undertaking the project with the lender's money are potentially unlimited but the losses are limited to the borrower's own resources invested in the project. Asymmetric information might result in adverse selection and moral hazard problems as

[34] For ease of argument, we draw Curve II parallel to Curve I, although the distance between the two could vary across different interest rates.

discussed above. Thus, limited liability and asymmetric information combine to create incentives for borrowers to undertake riskier projects. Given that limited liability truncates the downside losses to the borrower, the expected return to the borrower can be higher than the expected return of the project itself. Therefore, in a world of agency problems, a borrower has incentives to finance her project at interest rates above the project's expected return.

Curve III in Figure 4 is the demand curve that results from subtracting the demand-reducing effects of self-exclusion from Curve I and then adding the demand-increasing effects of limited liability and agency problems. We assume that the latter effects outweigh the former and, hence, we draw Curve III to the right of Curve I. Compared to Curve I, Curve III implies more loan applicants for any level of the lending interest rate (or the lending spread over $i_{mc}$). The distance between the two curves represents the severity of agency problems, as applicants in Curve III demand more external resources because they do not have to treat them as their own and are therefore willing to take greater risks and pay an interest rate higher than the expected return of their project. As in Figure 3, where the gap between interest rate and return to lender increases with increasing interest rates, it is likely that the distance between Curve I and III will increase with higher interest rates, as agency problems intensify.

Figure 4 allows us to define the universe of loan applicants or potential borrowers. For a given level of the lending interest rate, the corresponding point in Curve III gives the maximum number of loan applicants that would be observed, given the state variables, in a world characterized by self-exclusion, limited liability, and agency problems. The set of loan applicants grows bigger as the lending interest rate declines and it reaches a maximum at point A, the intersection of Curve III and the parallel to the horizontal axis at $i_{mc}$. The maximum number of applicants willing to take loans at $i = i_{mc}$ is denoted by $B_{III}$. The marginal applicant in

$B_{III}$ has the lowest expected return which, due to agency problems, is higher than the expected return of the project itself. Further, if $i_{mc}$ decreased to $i_{mc}$' (as shown in Figure 4) due to an improvement in the macroeconomic state variables, the maximum set of applicants would increase to $B_{III}$'. Thus, a reduction in systemic risk has a direct impact on total loan demand. Note, however, that along Curve III there will be an imprudently high demand for loans, an excessive number of potential borrowers compared to the number of applicants that what would obtain under Curve I, i.e., in the absence of the incentives distortions introduced by agency problems.

Let us now turn to the supply schedule illustrated in Figure 5, which is again drawn for a given type of loan and where the marginal lending interest rate is measured in the vertical axis. However, as discussed earlier, there is not a unique combination of costs, expected return, and risk that maps one-to-one to the interest rate. Rather, it is a mix of different borrower characteristics as perceived by the lender that determines the marginal borrower at each interest rate along the supply curve. To capture supply-driven changes in access, while keeping some indirect reference to demand, we follow a strategy proposed in De la Torre and Schmukler (2006a), namely, to focus on the *share* of loan applicants at a given interest rate that actually receive a loan. To determine such share, we use Curve III in Figure 4.[35] Thus, the horizontal axis in Figure 5 denotes the *share* of loan applicants receiving loans, which obviously reaches a maximum at 100%, denoted by the vertical line D. Each point along the this vertical line corresponds to a point in Curve III of Figure 4; as the interest rate increases along the vertical line D, the number of borrowers demanding loans falls (as per Curve III), while the share remains of course constant at 100%. Line D is thus a "normalized" demand curve, where the

[35] De la Torre and Schmukler (2006a), by contrast, discuss changes in access with reference to the demand for loans for "viable projects," which coincides with our definition of demand Curve I in Figure 4.

underlying level of demand, i.e. number of loan applicants, falls as the marginal interest rate increases, but the remaining loan applicants are re-normalized to one. Line D is truncated at $i_{III}$, the lending interest rate at which Curve III intersects the vertical axis in Figure 4 and, hence, where the pool of loan applicants is empty.

It is the lender who determines which borrowers get a loan. At each lending interest rate, the pool of loan applicants that request loans from the lender is determined by Curve III in Figure 4. It is a pool that includes "good" borrowers, who would treat the lender's money as if it was their own, as well as "bad" borrowers, who would not invest their own resources in projects for which they now demand a loan but would be willing to gamble with the loan resources if and once obtained. Faced with this problem, the lender uses a mixture of price and non-price devices to try to select the "right" debtors out of this pool. Her ultimate objective is to identify the subset of loan applicants that corresponds to Curve I in Figure 4, that is, to sort out those loan applicants that would treat the loan resources as if they were their own. The supply curve in Figure 5 is upward sloping but becomes steeper and steeper and eventually can bend backward. At $i = i_{mc}$ the share of loans applicants that get a loan from lenders is zero, as lenders would rather just invest in debt securities of the local government. The shape of the supply curve reflects a complex interaction between lenders' costs and their ability to identify appropriately expected returns and risks. Let us explain this better.

While the number of potential borrowers in the pool of loan applicants falls with increases in the lending interest rate (along demand Curve III in Figure 4), the riskiness of such pool increases. Nonetheless, in the lower region of the supply curve, lenders provide loans to a rising share of loan applicants as the interest rate increases. This is not just because a higher lending spread enables lenders to cover more costs. It is also because lenders feel confident in

their ability to assess risks and thus to sort out appropriately such debtors/projects that have higher expected risk-adjusted returns—and that would be thus able to repay the loan under reasonable states of the world. However, as the lending interest rate increases further, the task for the lender becomes increasingly more difficult because the pool of loan applicants becomes riskier (and thus the distance between Curve III and Curve I in Figure 4 widens). The lender becomes more and more reluctant providing loans to a larger share of loan applicants at higher interest rates (the supply curve steeps up) and, at a threshold interest rate, the supply curve reaches a flexion point and may begin to bend backward. *Prudent* loan suppliers thus try to offset the negative effects of higher interest rates on default probabilities via rationing—denying loans to debtors perceived to be too risky. While rationing intensifies as the lending interest rate rises, some rationing can happen at any point in the supply curve, in the sense that the share of borrowers that actually get a loan might not include some or many of the viable borrowers in Curve I of Figure 4. This is an inevitable cost of prudence in a world of uncertainty and agency problems—some "good" debtors *have* to be left out in order to minimize the set of "bad" debtors that are let in.[36]

Curve S* in Figure 5 denotes the potential supply curve for loans, i.e., the share of borrowers/projects that can be efficiently and prudently serviced by lenders as a function of the marginal interest rate and given the state variables, including $i_{mc}$. Prudent lenders that maximize risk-adjusted profits will supply loans up to the flexion point of S*, at i*, where the contractual income gains of a higher interest rate are just offset by the expected default losses. Therefore, i* is the "rationed equilibrium" marginal interest rate in the market for a given type of lending

[36] This is the classic tradeoff between "type I" and "type II" errors. The fuzziness in screening introduced by uncertainty and agency problems confronts the lender with overlapping probability distributions of good and bad debtors. Faced with this problem, lenders will in general try to minimize the error of letting too many bad debtors in, which necessarily implies that some good debtors will be left out.

product, and the flexion of S* at point I denotes the Access Possibilities Frontier for credit services, i.e., the maximum equilibrium outreach in terms of access to credit that is *prudently* achievable given the state variables, including $i_{mc}$. This corresponds to the bankable *share* of loan applicants A. Note that this bankable share depends on the level of i*, as the pool of loan applicants is different for different levels of the lending interest rate (as per Figure 4). 1 - A is the unbankable share of loan applicants at i*, given the state variables.[37]

Again, we allow for the possibility that the financial system may not achieve this potential so that actual supply, S, falls short of S*. Thus, a sub-optimal credit supply can obtain, for instance, if creditors' profits are not threatened by the pressures of competition and, as a result, they have little incentive to undertake greater but feasible cost and risk management efforts to reach out to new (costlier and riskier) clients.

We can use Figures 4 and 5 to identify different types of access to finance problems. To this end, we focus on deserving debtors/projects that do not get loans at the "rationed equilibrium" marginal interest rate, rather than on the number of deserving projects.[38] We start by identifying access problems that arise holding the state variables constant. A first type of access problem in this context would consist of too low a number of loan applicants simply because of self-exclusion resulting from cultural barriers or financial illiteracy. This would be reflected explicitly in a lower demand curve (Curve II) in Figure 4 but would only be implicit in the vertical line D in Figure 5.

A second type of access problem can arise due to supply sub-optimization, represented in Figure 5 by the supply curve S, which is below the curve S*. In this case, the problem is that of

[37] To get a measure of the overall unbankable share of loan applicants we would have to add loan applicants that are rejected at lower interest rates but decided not to demand a loan at this higher interest rate i*.

[38] Note that a key problem in emerging markets could be a lack of investment projects that deserve financing based on their expected return. This would be reflected in demand curve in Figure 4 that is too far to the left or too steep. While this is a relevant problem, it is not an "access to finance problem."

too low a supply outcome where, due to, for instance, regulatory distortions or insufficient contestability, lenders do not fully exploit all the outreach opportunities, given the market interest rate, the available loan technologies, and the state variables. As drawn in Figure 5, this not only results in a higher maximum marginal interest rate, but also a lower banked share of loan applicants at interest rate i*; the distance A - B at i* is a measure of access problems driven by supply inefficiencies.[39]

A third and very different access problem is associated with "excess access", denoted by supply to the right of the *prudent* possibilities frontier given at point I in S*. Curve S', to the right of S*, would in effect imply that loans are granted to a larger share of loan applicants than is prudently warranted, given the lending interest rate, and given the state variables, particularly the quality of the informational and contractual environment. For example, at i*, the problem of "excess access" can be represented by the distance C – A in Figure 5, which denotes the increase in the share of loan applicants receiving a loan at i* that is not warranted by prudent risk management. Point C in Figure 5 would be, given i*, "too close" to the vertical line D. "Too close" in the sense that lenders would be allowing in too many borrowers that are in Curve III but are not in the ideal Curve I (both in Figure 4).

While the points of demand-driven self-exclusion, supply-driven constrained sub-optimality and imprudent excessive access are defined holding constant state variables, there are two other access problems that refer directly to the location of the prudent access possibilities frontier and thus to deficiencies in state variables compared to countries with similar levels of economic development. The first such access problem would be due to an $i_{mc,}$ the opportunity

[39] There are two effects explaining the difference between interest rates at the flexion points of S and S*. On the one hand, S reflects higher costs resulting in higher interest rate for every borrower; on the other hand, the flexion point obtains earlier, so that the "equilibrium" marginal interest rate for S could also be lower than for S*. In the context of Figure5, however, we conjecture that the first effect is stronger than the second.

costs of lending, which is too high. The effect of lower opportunity costs is illustrated in Figures 4 and 6. A lower $i_{mc}$ will increase the universe of potential loan demanders (from $B_{III}$ to $B_{III}$' in Figure 4). It will also give rise to supply curve S*', which is to the right of S*, with a higher share of loan applicants served at any interest rate. For example, at i = i*, the share of loan applicants receiving finance would increase from A to B in Figure 6.

A second access problem related to state variables would arise due to a contractual and informational environment that is too weak compared to that in countries with similar levels of per capital income and which would affect the curvature of S*. An improvement in the contractual and informational environment would result in a supply curve that is flatter than S*, allowing for a larger share of loan applicants to be prudently and efficiently served by lenders at all relevant levels of the interest rate. The effect of an improvement in the state variables is illustrated in Figure 6 by curve S*''; the effect on the bankable share of loan applicants at i* would be the horizontal distance A-C. [40]

***c) Policies to Foster Access to Credit Services***

In order to design effective policies to expand access, policymakers have to be aware of whether their financial system is inside the frontier, at the frontier, or unsustainably beyond it, and/or whether the possibilities frontier is too low relative to countries of comparable levels of economic development. As in the case of payment and savings services, we can distinguish between *market-developing* policies—which raise the sustainable prudent possibilities frontier by changing the state variables—and *market-enabling* policies—which provide incentives to, or

[40] As in the case of the flexion points of S and S* in Figure 5, there are two offsetting effects on the flexion points as we move from S* to S*'. On the one hand, a lower $i_{mc}$ implies a lower interest rate for every borrower; on the other hand, this allows lenders to increase the marginal interest rate more. For ease of discussion, we assume that both effects cancel out in Figure 6, so that the flexion points of both supply curves are at the same interest rate i*.

remove obstacles for, private financial institutions to move closer to the frontier, given the state variables. Additionally, however, we have to consider *market-harnessing* policies that prevent the financial system from moving to an unsustainable imprudent equilibrium beyond the frontier. The weight given to each type of policies will depend on the diagnosis of the access to finance problems and an identification of where the more binding constraints lie.[41] It will also depend on the policy time horizon.

Consider first *market-developing policies*. As they aim at improving the state variables, these policies involve fundamental reforms that often confront political resistance and are thus difficult to implement. Market-developing reforms include, for instance, legal and even constitutional changes, major modernization processes in public sector agencies (including civil service reform), and substantial upgrades in macroeconomic, particularly fiscal, performance. To be sure, the results of market-developing reforms can at times be elusive—not least because of path dependence and problems associated with partial reform of complex institutional arrangements.[42] However, their pursuit must be given high priority where the binding constraints consist of major deficiencies in the state variables. Market-developing policies can be categorized according to the state variable they refer to: macro volatility, contractual and informational environment, or the scope for idiosyncratic risk diversification. Let us briefly consider each of these cases.

A high opportunity cost of lending may in some cases be the most important hindrance to the broadening of access to credit to the private sector. It may reflect doubts about fiscal solvency and a history of asset confiscation or inflation volatility, and will be often associated

[41] This very much resembles the "Growth Diagnostics" analysis by Hausmann, Rodrik and Velasco (2005).

[42] See De la Torre and Schmukler (2006b) for a discussion on the elusiveness of results of capital markets reforms in Latin America.

with high financial crowding out—i.e., the absorption by the government of a large share of society's financial savings. Under such circumstances, the appropriate policies to foster access to credit would be those that aim at enhancing the resiliency of fiscal solvency and at establishing a credible record of low and stable inflation.

A shallow credit market with low access may also obtain if there are major shortfalls in the contractual and informational frameworks.[43] The appropriate policies in this case would span a wide range: from titling of land property to the upgrading of laws affecting collateral repossession or execution of guarantees; from the modernization of corporate reorganization and bankruptcy proceedings to improvements in the functioning of the judiciary; from raising accounting and disclosure standards to establishing the appropriate legal framework and right incentives for the development of debtor information systems. Improvements in the contractual and informational environment facilitate the expansion of arm's-length financing: equity and debt contracts that are more impersonal in nature and that, therefore, rely more on transparency and enforcement rules of general application.

A final class of market-developing policies aims at raising the possibilities frontier by widening the scope for idiosyncratic risk diversification and thus, essentially, completing markets. To a significant extent, policies in this category coincide with those that aim at fostering macroeconomic stability. A minimum of macro stability, combined with sound management of the public debt and actions to the develop the local currency debt market are key in the formation of a reliable yield curve, which is in turn crucial to the development of interest

[43] There is considerable empirical evidence showing the adverse effects of a weak contractual and informational environment on credit depth and access. See Beck and Levine (2005) for an overview and IDB (2005) for an extensive survey of issues regarding creditor rights and debtor information systems in Latin America, providing ample evidence of the adverse role of poor creditor rights systems on credit depth, volatility, and access. There is also considerable empirical evidence showing that smaller, younger and more opaque firms are more financially constrained in economies with less developed financial and legal institutions and with less effective systems of credit information sharing (Beck, Demirguc-Kunt, and Maksimovic, 2005; Beck, Demirguc-Kunt, Laeven and Maksimovic, 2006; Love and Mylenko, 2003).

rate derivatives. Similarly, macro stability and a flexible exchange rate regime are necessary in the development of the market for currency derivatives. In small and undiversified economies, however, these policies might not suffice to give rise to markets for hedges and certain insurance products. In such cases, a premium should be placed on accessing the vast, yet vastly unexploited, risk pooling and diversification opportunities offered by international capital markets—which range potentially from catastrophic insurance to commodity price hedges, from weather and crop insurance to securitization of export receivables, and from currency swaps to GDP-indexed securities.[44]

Changes in the state variables involve changes in fundamental institutions and, thus, take a long time to materialize. That does not necessarily mean, however, that that there is no scope for policy in the shorter-run. To the extent that a financial system is operating below the possibilities frontier, there is room for *market-enabling policies* that may foster the broadening of access even in the absence of perceptible changes in state variables. Where the main reason for being below the possibilities frontier is the demand problem of self-exclusion, the appropriate policies would emphasize raising financial literacy. If – as is more likely - the main problems reside with sub-optimization in credit supply, by contrast, a wider range of policy options can be considered, starting with competition policy.

Competition is one obvious area for market-enabling policies aimed at enhancing supply outreach. Greater competition can result in efficiency gains that would move the system closer to the possibilities frontier. The salutary effects of greater competition are illustrated, for instance, by the recent vigorous expansion of sustainable and profitable micro- and consumer

[44] For an example of weather insurance, see Hess (2003). See OECD (2005) for a collection of papers on catastrophe insurance and Auffret (2003) for a discussion of the catastrophic insurance market in the Caribbean. Schiller (2003) provides a grand vision of the potential for financial markets to aid in long-term and international risk management, including reducing "gratuitous inequality", the feasibility of which is within reach given recent development in financial theory, information technology, and the science of psychology.

lending in emerging markets. It was spurred in many cases by the removal of unnecessary regulatory distortions to competition—including, for instance, allowing the use of scoring methods in credit risk management, calibrating entry capital requirements for microfinance institutions, abolishing of interest rate caps and providing regulatory incentives to the use of credit bureaus in screening debtors.[45] As competition thrived in the microcredit market, it stimulated innovation in loan products and the maximum use of available technologies, with consequent improvements in cost and risk management. The result has been a sustained broadening of access to micro credit, even where state variables have changed little.[46]

Beyond targeting competition per se, market-enabling policies can also try produce a movement towards the possibilities frontier by addressing hindrances such as coordination failures, first mover disincentives, and obstacles to risk distribution and sharing. While not easy to define in general terms, given their variety, these government interventions tend to share a common feature in creating incentives for private lenders and investors to step in, without unduly

[45] In its early stages of the development, microfinance tended to mitigate agency problems through group monitoring, i.e., lending to groups of borrowers rather than individuals with each member liable for the group loan. With the maturation of the microfinance industry, individual lending has become dominant. As regards collateral, rather than taking traditional collateral, microfinance institutions often take security on crucial consumer durables such as TVs and refrigerators, since these have a higher personal value for borrowers and thus reduce ex-post moral hazard problems. See Armendariz de Aghion and Morduch (2005) for an overview.

[46] Christen (2000) discusses its evolution from an NGO-based grant-funded activity to a vigorous and self-sustaining industry and associated "mission drift." Otero and Marulanda (2005) examine mayor trends and features in microfinance in Latin America, noting the simultaneous phenomenon of "upscaling" (the transformation of NGOs into regulated entities) and "downscaling" (the entering of well-established commercial banks into the microfinance business). The current status of microfinance and the opportunities and challenges going forward are discussed in World Bank (2006). Microfinance case studies include Navajas, Schreiner, Meyer, Gonzalez-Vega, and Rodriguez-Meza (2002) for Bolivia; Benavente (2006) for Chile, and Zaman (2004) for Bangladesh.

shifting risks and costs to the government.[47] For all the questions these market promoting interventions raise—and their systematic study is at an early stage to provide definitive answers—they clearly stand in sharp contrast with the traditional and typically ill-fated *market-substituting policies* consisting of government-directed lending at subsidized interest rates or through government-owned institutions.

Given the central role of risk in credit services, there is a third category of government policies related to access, which we define as m*arket-harnessing* and which try to prevent the financial system from moving to an unsustainable equilibrium beyond the frontier due to imprudent lending. Such imprudent lending binges can arise from the same competition that market-enabling policies try to foster if not accompanied by a proper defined regulatory and supervisory safety net. Indiscriminate free entry for new deposit-taking credit institutions or the intensification of competition among existing institutions can be a recipe for a banking crisis in waiting, in the context of implicit or explicit government guarantees, poor accounting and disclosure practices, deficient early warning system and prompt corrective action regimes and dysfunctional failure resolution frameworks.[48] Market-harnessing policies therefore aim at keeping banks' incentives to take aggressive risks in check through a mix of measures aimed at

[47] An analysis of this type of "market-friendly roles for the visible hand" is found in De la Torre and Schmukler (2006a), which presents case studies of such intriguing examples as: (i) the creation by NAFIN (a Mexican development bank) of an internet-based market for the discounting of post-delivery receivables by SMEs; (ii) a Chilean program (FOGAPE) to promote lending to SMEs via the auctioning of partial government guarantees; and (iii) a variety of structured finance packages orchestrated by FIRA (a Mexican development fund) to finance agricultural production (e.g., shrimp, corn); these packages succeed in involving, through risk sharing arrangements and partial guarantees, key stakeholders (multinational commercialization firm, suppliers, producers, banks, etc.) so to generate finance that would have otherwise not materialized.

[48] See Carletti and Hartmann (2003) for an overview over the theoretical and empirical literature on stability and competition. De Juan (2002) has a classical discussion of how, under conditions of unsound competition, good bankers turn into bad bankers. See also De la Torre et al. (2001) and De Krivoy (2000) for a discussion of how the weak application of fit-and-proper criteria at entry as well as deficient official monitoring led to unhealthy competition and banking crises in the cases of Ecuador and Venezuela, respectably. See Gavin and Hausmann (1996) for a discussion that even in the absence of moral hazard and deficient supervision competition can lead to crisis, as "good times are bad times for learning," in the sense that information asymmetries intensify during credit booms, which often result in crises.

strengthening market and supervisory discipline. The absence of too generous deposit insurance – implicit or explicit – and disclosure and transparency requirements give large creditors and depositors incentives and possibilities to monitor and discipline banks. In addition, market signals in the form of deposit interest rates, yields on subordinated debt or equity prices of publicly listed banks moving in response to risk taking and performance of banks provide additional information to bank supervisors and should be coupled with effective official intervention into institutions that the market has identified as weak. Market-harnessing policies are also important on the demand side to avoid predatory lending, which results in unsustainable overborrowing by individual borrowers. Predatory lending practices flourish in environments where borrowers face information gaps, imperfect competition and inadequate legal consumer protection. Disclosure requirements and education programs can help reign in loan sharks, while usury laws are much more difficult to calibrate.[49] In all, market-enabling policies focused on enhancing competition must ensure that gains in efficiency and access have to be coupled with market-harnessing policies that keep banks' incentives to take aggressive risks in check.

[49] See Honohan (2004c) for a more detailed discussion.

## IV. Concluding Remarks

This paper introduced a new tool to evaluate the outreach of a country's financial system and help design policies to increase prudent access and outreach. On the supply side we have focused on the fixed component of transaction costs and on lending risk as barriers for the financial system to reach out to clients with demand for low-value payment and savings transactions and to riskier borrowers. We distinguished between potential supply denoting the maximum outreach a financial system can provide given state variables such as the contractual and informational frameworks, the macroeconomic environment, technology and other country characteristics, and actual supply, which takes into account market structure and contestability resulting in no financial institution working at the frontier and catering to marginal customers. On the demand side, we distinguished between potential demand as predicted by economic factors and actual demand, which takes into account voluntary self-exclusion resulting in limited or no use of financial services by some customers. In the case of payment and savings services we defined the access possibilities frontier as the intersection of potential demand and supply and thus the bankable population as the share of population that could be served by the financial system given constraints imposed by the state variables and demand as defined by economic factors. Starting from this frontier, we defined three different access problems: first the lack of demand due to voluntary self-exclusion; second, supply of financial services below the potential due to lack of competition or other supply side constraints and third a frontier that is too low in international comparisons and explained by the state variables. Since credit risk is central to access to lending services, there is an additional possibility of excessive access due to imprudent lending practices.

The Access Possibilities Frontier has important implications for the debate on how to

expand access to financial services. It shows the potential for private solutions to expand access to financial services, but also points to an important role of the government in fostering such private solutions through competition and regulatory policies and through interventions characterized as private-public solutions. However, it also shows the limit to these efforts due to systemic deficiencies in the contractual, macroeconomic and informational environment that prevent the financial system from expanding access beyond a very low frontier. Finally, in the case of lending services, we discussed the role of government in harnessing market forces to avoid excessive imprudent access to credit.

The concept of the Access Possibilities Frontier is also useful for empirical attempts to measure outreach and access, bankable and banked population, for specific financial products.[50] A combination of household and supplier surveys can help compute demand and supply side factors and constraints, as we want to illustrate with the following example for the case of savings and payment services. First, household surveys can be used to determine the share of the population that uses a specific service, such as checking account, savings account, or ATM services (Point III in Graph 2). Second, using survey questions on the reasons why certain people do not use a specific service, one can identify the (i) unbankable (no sufficient income or economic activity), (ii) unbanked due to voluntary exclusion (by choice) and (iii) unbanked due to involuntary exclusion (access related obstacles such as costs, lack of necessary documentation, distance etc.). Adding the banked population, the unbanked by choice and the unbanked due to involuntary exclusion, one can get to a first approximation to the bankable population (point I in Graph 2) or the frontier. The relative shares of unbanked due to voluntary and involuntary

[50] Porteous (2005) provides an example for such a calculation for the Mzansi account in South Africa.

exclusion also allow quantifying the effect to be expected by market-enabling policies aimed at the demand and supply side.[51]

The tool of access possibilities frontier can also be used to assess the impact of cost- or risk-reducing technological change and of changes in state variables such as the legal system on the bankable population. The asset holdings of enterprises and households can be compared to collateral requirements of financial institutions and as impacted by the contractual framework. If legal system reforms allow a "new" asset (e.g. movable assets) to be used as collateral, enterprise and household surveys can be used to calculate the share of potential borrowers that would become bankable as they are now able to use their assets as collateral, thus advising policy makers on the binding constraint and most beneficial reform in terms of higher outreach.

While we have developed and discussed the access possibilities frontier on a conceptual level, we hope to have convinced the reader that it can very well be operationalized for both empirical and policy work. Disentangling different access problems and the impact of policy reforms is much more difficult for lending services than for payment and savings services, as the assessment and management of risk adds a crucial additional element. Nevertheless, we are confident that this tool can be refined further to serve both the debate on the measurement of access to finance as well as the policy debate on how to improve outreach of the financial system and access to financial services.

[51] Alternatively, the bankable population could be computed through estimates of the costs of delivering certain payment services, affordability assumptions (x% if monthly income) and information on the income distribution in the country. These estimates can then be compared to a supplier survey on the actual prices for the same products to calculate the share of the unbanked population that is excluded due to supply constraints. The notorious unreliability of such cost estimates, due to the important role of overhead costs in financial institutions, is an important caveat for such an exercise.

**References**


Armendariz de Aghion, B. and Morduch, J. (2005): The Economics of Microfinance, MIT Press, Cambridge, MA.

Auffret, P. (2003): Catastrophe Insurance Market in the Caribbean Region. Market Failures and Recommendations for Public Sector Interventions, World Bank Policy Research Working Paper 2963.

Beck, T., Levine, R. and Loayza (2000): Finance and the Sources of Growth, *Journal of Financial Economics 58*, 261-300.

Beck, T., Demirguc-Kunt, A., and Levine, R. (2004): Finance, Inequality and Poverty Alleviation: Cross-Country Evidence, World Bank Policy Research Working Paper

Beck, T., Demirguc-Kunt, A., and Maksimovic, V. (2004): Bank Competition, Financing Constraints, and Access to Credit, *Journal of Money, Banking and Credit* 36, 627-48.

Beck, T., Demirguc-Kunt, A., and Maksimovic, V. (2005): Financial and Legal Constraints to Firm Growth: Does Firm Size Matter?" *Journal of Finance* 60, 137-177.

Beck, Thorsten, Aslı Demirgüç-Kunt, Luc Laeven and Vojislav Maksimovic. 2006. "The Determinants of Financing Obstacles." *Journal of International Money and Finance*, forthcoming.

Beck, T., Demirguc-Kunt, A., and Martinez Peria, M. (2005); Reaching Out: Access to and Use of Banking Services across Countries, World Bank Policy Research Working Paper.

Beck, T., and Levine, R. (2005): Legal Institutions and Financial Development. In: Menard, C. and Shirley, M. (eds.), Handbook of New Institutional Economics. Kluwer Dordrecht.

Benavente (2006): Programa de Crédito para la Microempresa. Background paper for De la Torre and Schmukler (2006a): Innovative Experiences in Access to Finance: Market Friendly Roles for the Visible Hand, Latin America Regional Study. World Bank.

Bossone, B., Honohan, P and Long, M. (2002); Policy for Small Financial Systems, in: Caprio, G., Honohan, P and Vittas, D. (Eds.): Financial Sector Policy for Developing Countries: A Reader, World Bank. Washington, DC

Carletti, E. and Hartmann, P. (2003): Competition and Stability: What's Special about Banking?, in Mizen, Paul (ed.): Monetary History, Exchange Rates and Financial Markets: Essays in Honor of Charles Goodhart, Edward Elgar.

Christen, R.P. (2000): Commercialization and Market Drift: The Transformation of Microfinance in Latin America, CGAP Occasional Paper 5.

Claessens, S., Glaesner, T. and Klingebiel, D. (2001): Electronic Finance: Reshaping the Financial Landscape around the World, *Journal of Financial Services Research* 22, 29-61.

Claessens, S., Dobos, G., Klingebile, D. And Laeven, L. (2003): The Growing Importance of Networks in Finance and its Effects of Competition, in: Nagurney, A. (Ed): Innovations in Financial and Economic Networks, Edward Elgar Publishers, Northampton, MA, USA, 110-135.

Claessens, S. and Laeven, Luc (2004): What Drives Banking Competition? Some International Evidence, *Journal of Money, Banking and Credit* 36, 563-83.

Claessens, S. (2005): Access to Financial Services: A review of the Issues and Public Policy Objectives, World Bank Research Observer.

De Juan, A.(2002): From Good Bankers to Bad Bankers, in: Caprio, G., Honohan, P and Vittas, D. (Eds.): Financial Sector Policy for Developing Countries: A Reader, World Bank. Washington, DC

De Krivoy, R. (2000): Collapse: The Venezuelan Banking Crisis of 1994, Group of Thirty

De la Torre, A. Garcia-Saltos, R., and Mascaro, Y. (2002): Banking, Currency, and Debt Meltdown: Ecuador Crisis in the Late 1990s, World Bank mimeo.

De la Torre, A. and Schmukler, S. (2005): Coping with Risks Through Mismatches, *International Finance* 7, 1-42.

De la Torre, A. and Schmukler, S. (2006a): "Innovative Experiences in Access to Finance: Market Friendly Roles for the Visible Hand," Latin America Regional Study. World Bank.

De la Torre, A., Gozzi, J.C. and Schmukler, S. (2006b): "Capital Market Development. Whither Latin America?" World Bank.

Financial Services Authority (2000): In or Out? Financial Exclusion: A Literature and Research Review, Consumer Research 3.

Gavin, M. and Hausmann, R. (1996): The Roots of Banking Crises: The Macroeconomic Context, Inter-American Development Bank Working Paper 318.

Hausmann, R. Rodrik, D. and Velasco, A. (2005): Growth Diagnostics, Harvard University mimeo.

Hess, U. (2003): Innovative Financial Services for Rural India: Monsoon-Indexed Lending and Insurance for Smallholders. Agriculture and Rural Development Working Paper 9, World Bank.

Honohan, Patrick (2004a). Financial Development, Growth and Poverty: How Close Are the Links? In Charles Goodhart, ed. *Financial Development and Economic Growth*: Explaining the Links. London: Palgrave.

Honohan, P. (2004b): Measuring Microfinance access: Building on Existing Cross-country Data, World Bank Policy Research Working Paper 3606.

Honohan, P. (2004c): Financial Sector Policy and the Poor. Selected Findings and Issues, World Bank Working Paper 43.

Inter-American Development Bank (2004): Unlocking Credit. The Quest for Deep and Stable Bank Lending.

Karlan, D. and Zinman, J. (2004): Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment

Kempson, E., Atkinson, A. and Pilley, O (2004): Policy level Response to Financial Exclusion in Developed Economies: Lessons for Developing Countries, University of Bristol.

Levine, R. (1997): Financial Development and Economic Growth: Views and Agenda." *Journal of Economic Literature* 35, 688-726.

Levine, R. (2005): Finance and Growth: Theory and Evidence. In: Aghion, P and Durlauf, S. (eds.), Handbook of Economic Growth. Elsevier Science. The Netherlands.

Love I. and Mylenko, N.. (2003): Credit Reporting and Financing Constraints. World Bank mimeo.

Napier, M. (2005): Engaging the Private Sector: The Case for Financial Access Charters in Sub-Saharan Africa, Finmark Trust.

Navajas, F., Schreiner, M., Meyer, R.L. Gonzalez-Vega, C, and Rodriguez-Mesa, J. (2000): Microcredit and the Poorest of the Poor: Theory and Evidence from Bolivia, *World Development* 28, 333-46.

North, D. (1990): Institutions, Institutional Change, and Economic Performance, Cambridge, UK: Cambridge University Press.

Organization for Economic Cooperation and Development (2005): Catastrophic Risks and Insurance, Paris, France.

Otero, M. and Marulanda, B. (2005): The Profile of Microfinance in Latin America in 10 Years: Vision and Characteristics

Peachey, S. and Roe, A. (2004): Access to Finance: A Study for the World Savings Banks Institute.

Porteous, D. (2004): Making Financial Markets Work for the Poor, Finmark Trust

Porteous, D. (2005): The Access Frontier as a Tool in Making Markets Work for the Poor, mimeo.,

Rajan, R. and Zingales, L. (2003); Savings Capitalism from the Capitalists, Crown Business, New York.

Shiller, R.J. (2003): The new Financial Order: Risk in the 21st Century, Princeton University Press, Princeton, NJ.

Stiglitz, J. and Weiss, A. (1981): Credit Rationing in Markets with Imperfect Information, *American Economic Review* 71,393-410.

Williamson, S.D. (1987): Costly Monitoring, Loan Contracts and Equilibrium Credit Rationing, Quarterly Journal of Economics 102, 135-46.

World Bank (2004): India – Scaling-up Access to Finance for India's Rural Poor.

World Bank (2004): South Africa: Technology and Access to Financial Services, World Bank mimeo.

World Bank (2005): Access to Financial Services in Brazil, The World Bank.

World Bank (2006): Access for All: Building Inclusive Financial Systems.

Zaman, H. (2004): The Scaling-up of Microfinance in Bangladesh: Determinants, Impact and Lessons, World Bank Policy Research Working Paper 3398.

Figure 1: Scale economies in payment and savings services

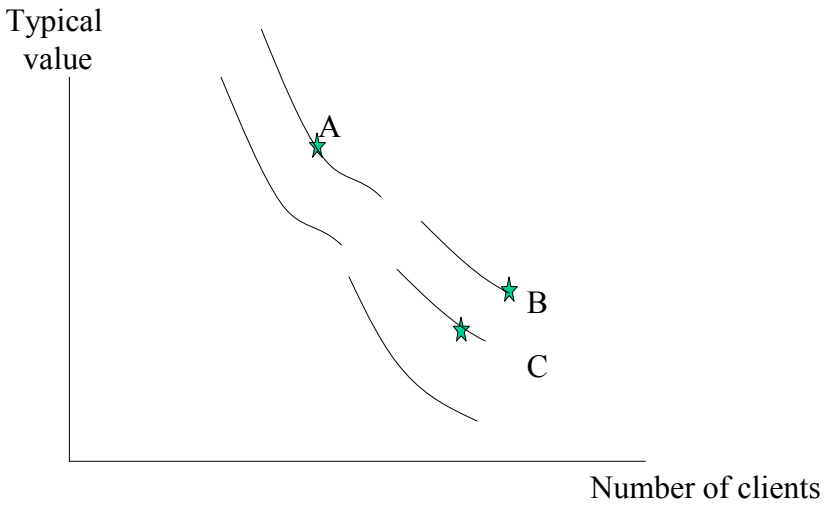


Figure 2: Access Possibilities Frontier for payment and savings services

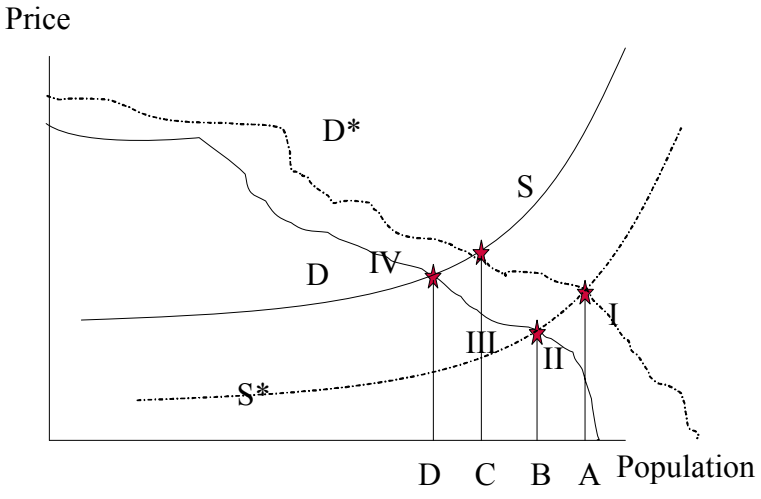

Figure 3: Systemic risk, agency problems, risk premiums and credit rationing

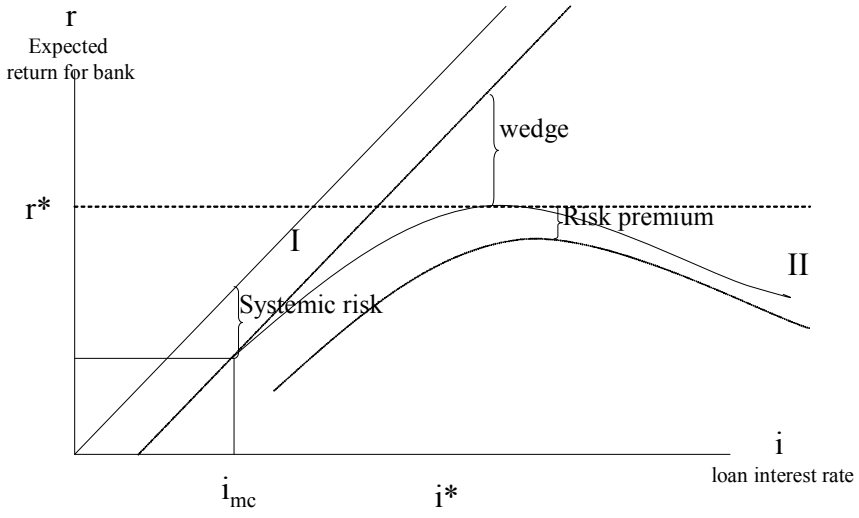


Figure 4: Loan demand as function of expected return, non-economic factors and agency costs

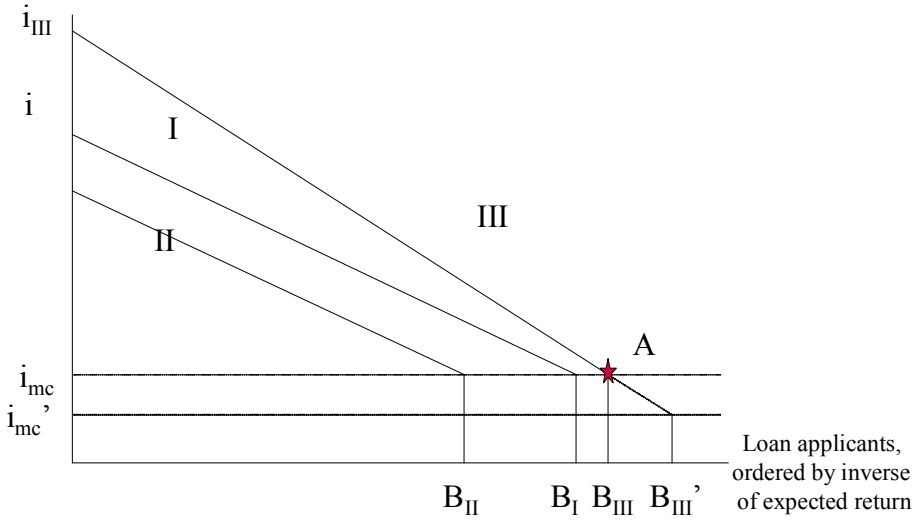

Figure 5: Access Possibilities Frontier for credit services

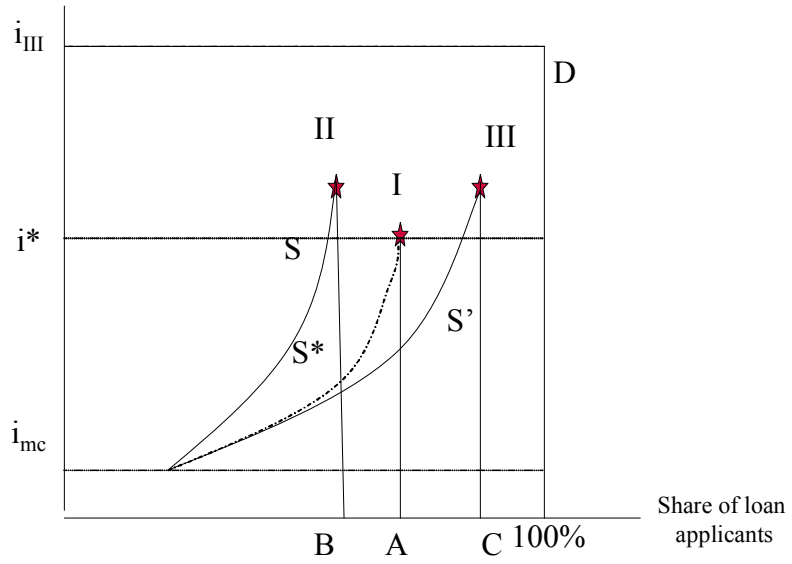


Figure 6: Access Possibilities Frontier for credit services – changes in state variables