Essays on Access to Financial Institutions, Inequality, and Redistribution

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Ulrike Vogelgesang

Vorgelegt im Wintersemester 2001/2002

Referent:	Prof. Axel Börsch-Supan, Ph.D.
Korreferent:	Prof. Martin F. Hellwig, Ph.D.
Dekan:	Prof. Dr. Christoph Buchheim
Tag der mündlichen Prüfung:	31. 01. 2002

Acknowledgements

This thesis has been prepared during my studies at the Graduiertenkolleg "Finanz- und Gütermärkte" at the University of Mannheim. It has greatly benefited from comments and encouragement from my colleagues, supervisors, and friends.

My supervisors Prof. Axel Börsch-Supan and Prof. Martin Hellwig substantially contributed to this work through their continuous advice and encouragement, helping me to keep focus.

I am grateful to Northwestern University for the hospitable research environment provided throughout my visit at MEDS in spring and summer 1999. Special thanks go to Prof. Benny Moldovanu who made this visit possible. My work on optimal capital income taxation has greatly benefited from discussions with Marco Bassetto, who helped me to connect my ideas with other strands of literature.

The second part of this thesis would not have been possible without the support from IPC GmbH, Frankfurt—especially Gabriel Schor—and the staff of Caja Los Andes, La Paz, Bolvia. Providing access to the data set from Caja Los Andes and giving me a unique opportunity to see "Microfinance in action" has greatly added to the fun and the satisfaction of my work.

Finally, I want to thank those friends that provided continuous support and helped through the harder times, most notably my colleagues from the GK, my "WG," Michèle, my family, and especially Ralf.

Stuttgart, April 2002

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Inequality in income and wealth has shown a high persistence over time. This thesis discusses different means to reduce inequality; firstly, through the use of redistributive taxation, secondly, through a new strategy to improve the opportunity set of poor households via better access to capital markets.

In 1999, 23% of all people in developing countries lived in extreme poverty, that is, on less than $US 1.08 a day (Worldbank 2000b). The richest fifth of the population obtained 61% of all income in Bolivia (1968), 44% in the USA (1991), and 39% in Germany (1984) (Deininger and Squire 1996). Besides an inequality in incomes, many countries have an even higher level of inequality in wealth, the latter being particularly persistent since wealth tends to be inherited from one generation to another.

Why should we worry about inequality? After all, inequality in incomes is a natural consequence of different abilities. People with a higher productivity can produce more than others and generate higher incomes as a result. In spite of this link, high inequality is met with widespread concern that is mainly driven by three aspects. Firstly, the existing levels of inequality are frequently perceived as excessively high, that is, larger than could be explained by different abilities alone. This is particularly so since the causes of inequality frequently seem to lie in differences in economic opportunities. Secondly, there is a growing body of evidence that high inequality is connected with frequent social unrest, high violence and crime rates (The Economist 2001a, Kelly 2000). Finally, democratically elected governments have an obligation to ensure the well-being of all individuals in their state and thus express a well-founded concern about increasing inequality.

Given that inequality is higher than we would like it to be, what can we do about it? The following chapters of this thesis discuss two possibilities to reduce inequality. Firstly: through the use of redistributive taxation (chapter 2), and secondly, through an improvement in the economic opportunities of low income households via a better access to capital markets (chapters 3 to 5). Each chapter stands alone and can be read as such; this introduction presents the common ground and discusses the underlying ideas.

On a nationwide level, the most direct means to reduce inequality is taxation. Why not tax the rich and pay transfers to the poor? While this idea has been appealing to numerous policy makers, economists have long pointed to the adverse incentive effects of high taxation (Ramsey 1927, for example). If income taxes are very high, people might work less than otherwise or increasingly work in areas outside the official tax system. If income taxes are progressive, people with the highest incomes (and, arguably, the highest productivity) have the least incentives to work an extra hour. One of the fundamental questions in this area then is how to balance the redistributive benefits of taxation and the adverse incentive effects. One particularly controversial area in this discussion has been the taxation of capital income. Due to the investing nature of savings, taxes on capital income not only decrease current income but potentially decrease future growth. This effect has resulted in overwhelming arguments in favor of zero capital income taxes, even when there is inequality. The first part of this thesis analyzes this question in more detail.

For an economy with arbitrarily many households, chapter 2 shows that if households are heterogeneous with respect to productivities and endowments, zero capital income taxes generally are not optimal. We find that zero capital income taxes can be optimal only if endowments are homogeneous, if the production function is weakly separable between labor and capital, and if utility functions are homothetic and identical across agents. In a simplified model we further show that the extent of the inequality and the joint distribution of its different components (productivities and endowments in our model) are crucial for the size of the marginal welfare effects of taxation. A positive correlation between endowments and productivities increases the marginal welfare effects of capital income taxation, while a negative correlation decreases the effects.

While redistribution constitutes a means to improve the situation of poor households, it addresses the consequences of inequality only and not its causes. What lies behind the high levels of inequality observed? One important factor is the inequality between rich and poor people in their access to institutions (The Economist 2001b). Institutions such as courts or the financial service sector provide the necessary support for functioning markets. The high barriers low income households face in accessing these institutions play a large part in the manifestation of poverty. One way to reduce inequality thus can consist in improving these households' access to institutions or to build new institutions where none exist. While the discussion above has focused on direct subsidies to poor households, one could also use these subsidies to improve poor households' economic opportunities instead. One crucial question thus is, given the same amount of money available, which strategy is more effective? The second part of the thesis focuses on a set of institutions that provide financial services to poor households and analyzes their effectiveness in increasing the incomes of their clients.

The missing access to financial markets is one of the most limiting factors for the economic opportunities of small entrepreneurs. Deprived from access to formal banks, many of these entrepreneurs pay horrendously high interest rates to moneylenders (Murinde 1996, Aleem 1993). To reduce the costs of loans for these entrepreneurs, the concept of microfinance was developed in the late 1970's and has become increasingly popular since. Compared to previous attempts to provide credit to poor households, the novelty of microfinance consists firstly in the use of new incentive mechanisms such as group loans or other collateral substitutes, and secondly in the attempt to cover costs through high interest rates which are a prerequisite for long-time, sustainable services (Krahnen and Schmidt 1994).

The second part of the thesis asks how these new institutions affect their clients and whether they can provide financial services on a sustainable basis. One important aspect of such an assessment lies in the analysis of the clients and their economic situation over time. Chapter 3 examines the development of the clients from one such institution and their enterprises and asks whether we can observe a lasting increase in the scale of their businesses and their incomes. For those clients who take out repeat loans and stay with the microfinance institution for some time, we find a strong increase in assets and business income. For clients in the commerce sector, for example, half of the clients who took their first loan in 1995 and were still clients of the microfinance institution in 2000, have increased their assets by more than 325% and their business incomes by more than 60% over these years.

For a thorough analysis of the contribution of these loans, it is necessary to address selection issues. That is, we need to ask how much of any observed increase in incomes can actually be attributed to the loans. Clients who obtain loans might simply be more productive than others and would have achieved a similar growth without those loans. Chapter 4 discusses the methodological concepts of an impact analysis and estimates the contribution of loans to growth in assets and to production efficiency. The results show a strong positive influence of loans on growth in assets. We also find that clients with prior loans generate higher sales revenues from the same amount of assets than clients without prior loans, indicating increased efficiency. Perhaps surprisingly, these effects are stronger for larger businesses.

In a final assessment of microfinance, chapter 5 analyzes repayment behavior. Microfinance can improve economic opportunities in the long term only if it works on a sustainable basis. A prerequisite for sustainability are high repayment rates. Chapter 5 analyzes repayment in a particularly tumultuous situation where a rising indebtedness of clients, an economic crisis, and increasing competition of microfinance institutions coincide with a pronounced increase in late payments and capital at risk. Our results suggest that the following factors contributed to rising arrears. Firstly, distributing more loans to clients who already have other loans leads to lower repayment rates. Secondly, clients with overdue payments in their prior loans are significantly more likely to pay late for future loans as well. This strong correlation suggests that capital at risk could be reduced by following a stricter policy in rejecting loans for clients with a bad repayment record. Thirdly, a tolerance of payments with a few days overdue leads not only to a higher probability that payments are late but also to a higher probability that they remain overdue for many days and add to capital at risk. With respect to the economic environment, we find that late payments are not driven by low economic growth. In addition, we find that clients with given characteristics are more likely to pay on time when there is high competition and a high supply of micro-loans than otherwise. We also find that clients are more likely to pay late in repeat loans than in their first loans and that women are more likely to pay late than men.

The results presented in this thesis show two successful means for the reduction of inequality. The first part demonstrates that it is desirable to tax capital income and to use the proceeds to redistribute unless inequality and production functions are of a very special form. The second part of the thesis discusses microfinance as a means to improve the opportunity set of poor households. We show that microfinance has the potential of increasing the incomes of poor households and that it can maintain sustainable repayment rates even in difficult times.

Part I

Fiscal Policy

Chapter 2

Optimal Capital Income Taxation and Redistribution*

2.1 Introduction

The study of optimal tax systems in a dynamic framework has mainly focused on efficiency aspects.[1] In the present paper we choose a different approach, focusing on the impact of agent heterogeneity on optimal tax rates, where taxes are collected for redistributional purposes. Under the assumption that the government maximizes a social welfare function, we ask whether non-zero capital or labor income taxes can be optimal for different sources of inequality. How do correlations between labor income and wealth interact with the welfare effects of taxation?

To answer these questions, we first review the literature on capital income taxation with heterogeneous agents and apply the results of the literature on uniform commodity taxation to a dynamic setting. We then develop a two period model, in which households are heterogeneous with respect to their endowments and abilities. Remaining tractable analytically, this model allows us to study the effects of different sources of heterogeneity among households on optimal tax rates. In particular, we show that the optimal capital income tax rate in general is non-zero. The welfare

*This chapter has been published in Finanzarchiv (2000), issue 57 (4), pp. 412-434, which generously has permitted the inclusion in this thesis.

[1]See, for example, Chamley (1986) and Jones, Manuelli, and Rossi (1997).

effects of taxation depend crucially on the joint distribution of initial wealth and productivities and on the flexibility of labor income taxes over time.

How do the results presented in this paper relate to earlier findings on optimal capital income taxation? Infinite horizon representative agent models in the tradition of Chamley (1986) show that, in the long run, it is not optimal to finance an exogenous stream of government expenditures through capital income taxes. However, there are a few initial periods (their number depending on a possible upper bound on tax rates) in which the optimal capital income tax is strictly positive, declining to zero afterwards. The main effect of initially high levels of capital income taxes is to extract the endowments from the consumers. The government builds a large surplus in the initial periods from which it finances part of its expenditures thereafter.

A limitation of this approach is its reliance on a representative agent with an infinite horizon. In such a setting, intragenerational distribution is not an issue. Another limitation of most infinite horizon models is the assumption that the government is allowed to build up a substantial surplus in the early periods, which is often limited only by the assumption that taxes ought to be no higher than 100%. Tax rates of this magnitude might be hard to implement.

Judd (1985) considers an infinite horizon model with two types of agents. In his most general setting, the agents differ with respect to their initial endowments and utility functions. Agents derive utility from consumption and leisure. The government has a fixed stream of expenditures over time and raises revenues through capital and labor income taxation. It redistributes income with a non-negative lump-sum transfer, which may be different for both types of agents. Production is weakly separable between capital and labor. Within this framework, Judd shows that if there exists a steady state, then in this steady state it is not optimal to tax capital income. Chari and Kehoe (1999) build a similar model without lump-sum transfers. Redistribution, thus, is a side-effect of revenue raising. They confirm Judd's results and show that the assumption of a weakly separable production function is necessary for the optimality of zero capital income taxes.

While these models show that the optimal capital income tax rate in the steady state is zero, they offer little insight about optimal rates off the steady state. Our model, in contrast, considers a finite number of periods and shows that zero taxes on capital income are not optimal if, for example, goods endowments are heterogeneous.

The models discussed above assume perfectly competitive markets. In addition to these, there are a number of studies analyzing optimal capital income taxes in the presence of market imperfections. With few exceptions, they find that the optimal capital income tax rate is different from zero. Judd (1997), for example, shows how monopolistic competition among firms can lead to the optimality of a negative tax on capital income. Aiyagari (1995) and Chamley (2001), on the other hand, find that incomplete credit markets can lead to the optimality of a positive capital income tax rate. We do not analyze such market imperfections but show that even with complete markets it can be optimal to impose a strictly positive tax on capital income.

The study of optimal capital income taxation is closely related to earlier work on uniform commodity taxation, such as Atkinson and Stiglitz (1976). Their static analysis can be reinterpreted as a dynamic model where different commodities represent consumption at different points in time. The relation between these approaches will be explored below. In addition, our model is related to the literature on optimal linear income taxation. Sheshinski (1972) shows that in a static setting with heterogeneity in the agents' productivities, the optimal marginal income tax rate is strictly positive and less than 100%. The model presented in this paper considers heterogeneity in two dimensions (in productivities and goods endowments) and shows under which conditions the one-dimensional result holds.

The remainder of the paper is organized as follows. Section 2.2 builds upon the existing literature on optimal capital income taxation and uniform commodity taxation, and discusses its implications for optimal taxation with heterogeneous agents. In section 2.3, we construct a simplified model emphasizing the interaction between the households' two-dimensional heterogeneity and marginal welfare effects. A conclusion is provided in section 2.4.

2.2 Optimal Taxation with Heterogeneous Agents

This section provides an overview of the effects of heterogeneity on optimal taxation. The existing literature provides little analysis of optimal capital income taxation with heterogeneous agents.[2] The few articles about this topic largely restrict themselves to steady state analyses. However, there is a body of literature based on Atkinson and Stiglitz (1976) that considers optimal taxation with multiple factors and heterogeneous agents. Atkinson and Stiglitz show that it is not optimal to distort relative prices of consumption goods if a sufficiently flexible income tax scheme is available. In a dynamic interpretation, their result implies that the optimal capital income tax rates are zero. Their results are relevant for our analysis since labor in different periods and endowments may be interpreted as different factors of production.[3]

In the following paragraphs we apply the results of this strand of literature to the issue of optimal capital income taxation and show under which conditions optimal capital income tax rates can be zero. We further show how and why the required sets of assumptions differ with time-separable utility and in the steady state.

2.2.1 Optimal Capital Income Taxation

The literature on commodity taxation shows under which conditions optimal commodity tax rates are uniform. Applying the results of this literature to a dynamic setting, we obtain our first proposition.

PROPOSITION 1 *Suppose there is an arbitrary number of agents who are heterogeneous with respect to their productivities and goods endowments. Utility is separable between leisure and consumption and is strictly concave in all arguments. There is an arbitrary number of goods. A social planner sets linear tax rates on labor and capital income and distributes lump-sum transfers to maximize welfare, which is de-*

[2] One exception is Chamley (2001). The focus of his paper, however, is on credit market constraints.

[3] For a detailed account of the relation between static and dynamic models of this kind see the first sections in Judd (1997).

fined as the weighted sum of all agents' utilities. A capital income tax rate of zero then fulfills the planner's necessary conditions if all of the following conditions hold.

i. Preferences for consumption are homothetic and identical for all households.

ii. Goods endowments are homogeneous or proportional to actual consumption.

iii. Production is either weakly separable between labor and capital[4] or agents have identical productivities.

If utility is weakly separable only, condition (i) requires that preferences for consumption and leisure are homothetic and identical.

Items (*i*) and (*ii*) correspond to the findings in Bassetto (1999).[5] In a setting similar to Atkinson and Stiglitz (1976), he analyzes properties of optimal commodity taxes for general homothetic and separable utility functions. His model considers two types of agents, of whom only one works, and he finds that homothetic and identical preferences as well as homogeneous or proportional endowments lead to the optimality of uniform commodity taxes or, in the dynamic interpretation, to a zero tax on capital income.

Item (*iii*) of Proposition 1 is analogous to the results of Naito (1999). Reexamining optimal commodity taxation in a setting close to Atkinson and Stiglitz (1976), he shows that the optimality of uniform commodity taxes is not robust against the introduction of production functions that are not weakly separable between labor and capital. For a formal derivation of Proposition 1, see appendix 2.5.1.1.

The following paragraphs discuss the individual conditions of Proposition 1 and show why each of them is required for the optimality of a zero tax on capital income.

[4] In the remainder of the paper, weak separability of a function $F(x_1, x_2, y)$ between x and y is defined as $\partial(\frac{F_{x_1}}{F_{x_2}})/\partial y = 0$.

[5] Item (*i*) also corresponds to earlier findings by Atkinson (1977) and Deaton (1979).

Homothetic and identical preferences If preferences are not homothetic, luxury goods may exist. That is, wealthier agents consume disproportionately more of the luxury goods than poorer agents. Thus, while labor income taxes are proportional to the agents' productivities, higher tax rates on the luxury good disproportionately tax the wealthy and, thus, provide a means to redistribute. In a dynamic interpretation this example translates as follows. If, for example, the desire for consumption in later periods increases with income, wealthier households save disproportionately more than poorer ones. A tax on capital income, therefore, disproportionately affects the wealthy. If preferences differ across households, a similar mechanism works.

Homothetic preferences are required for zero capital income taxes being optimal only if labor income taxes are restricted to be linear. Atkinson and Stiglitz (1976) show that for sufficiently flexible non-linear labor income taxes uniform commodity taxation can be optimal if preferences are weakly separable between consumption and leisure. Homotheticity is not required for this result since non-linear labor income taxes already provide a means to tax the wealthy disproportionately. Even if there are luxury goods, an additional tax on these will not improve welfare if labor income taxes follow an optimal disproportionate scheme.[6]

Homogeneous goods endowments Heterogeneous endowments undermine the optimality of uniform taxation since they lead to different intertemporal trading patterns among agents. If wealthier agents have higher capital holdings due to higher endowments, a taxation of capital income extracts high revenues from the wealthier agents which can be used for redistribution. These tax payments are directly related to differences in endowments that are generally not captured by the revenues of linear labor income taxes, which are proportional to productivities.

Weakly separable production The last requirement in Proposition 1 concerns the production side of the economy. To see why this assumption is important, con-

[6] Similar results have been derived by Cremer, Pestieau, and Rochet (1999). Introducing a two-dimensional heterogeneity in a static Mirrlees-style model with non-linear labor income taxation and two consumption goods, they show that uniform commodity taxation is not optimal.

sider the following example. Let the production function be such that rising capital usage in production implies the decrease of relative productivity of low productivity households. If this is the case, the government might want to discourage capital accumulation by taxing capital income in order to prevent a higher discrepancy in relative productivities. In other words, if production is not weakly separable between labor and capital, capital income taxes might influence relative productivities. As a consequence, non-zero capital income taxes can be optimal even if endowments are homogeneous and preferences are homothetic and identical across households.

2.2.2 Optimal Labor Income Taxation

In analogy to capital income taxation, we can show under which conditions optimal labor income taxes can be uniform. Since wage income is indirectly taxed by capital income over time, we consider present value labor income taxes.

PROPOSITION 2 *Consider the same setting as in Proposition 1. Uniform present value labor income taxation fulfills the planner's necessary conditions if all of the following conditions hold.*

i. Preferences for leisure are linearly homogeneous and identical for all households.

ii. Productivities are homogeneous.

If utility is weakly separable only, condition (i) requires that preferences for consumption and leisure are homothetic and identical.

The intuition behind the conditions is similar to the discussion in the previous section. Homothetic and identical preferences ensure that the relative labor supply in different periods does not vary across households with different wealth. Thus, varying labor income taxes over time would not tax wealthy households disproportionately. If productivities are heterogeneous, homothetic preferences are not sufficient to ensure a proportional labor supply for all households. Homogeneity of endowments, however, is irrelevant for the optimality of uniform labor income taxation. For a formal analysis see appendix 2.5.1.2.

2.2.3 Time Separable Utility and Optimal Taxation in the Steady State

While the results presented in Proposition 1 are derived from a static setting with multiple goods, dynamic analyses virtually always assume time separable utility functions. How does this assumption modify the requirements for the optimality of a zero tax on capital income as presented in Proposition 1? Most work on optimal capital income taxation is further restricted to a steady state analysis. In a steady state, consumption and labor are constant over time and there are no endowments. How do these steady state-assumptions modify the above conditions?

PROPOSITION 3 *Consider the same environment as in Proposition 1 and assume utility is time-separable. A capital income tax rate of zero fulfills the planner's necessary conditions if all of the following conditions hold.*

i. Preferences for consumption are homothetic and identical for all households.

ii. Goods endowments in periods t and $t-1$ are homogeneous or proportional to actual consumption.

iii. Production is weakly separable between labor and capital or agents have identical productivities.

In a steady state with no endowments and constant consumption and labor, item (iii) is sufficient.

While homotheticity and weakly separable production are required for the same reasons as before, time separability limits the effects of heterogeneous endowments to two periods. Capital income taxes in the period with heterogeneous endowments are used to redistribute while next period's capital income taxes ensure that relative prices in all following periods are not affected. Longer lasting effects occur only if the capital income tax rate is restricted, for example, to be no larger than 100%.[7]

[7] This intuition lies behind the result in Chamley (2001) who finds that the optimal capital income tax rate is zero in finite time, even if endowments are heterogeneous.

The mechanisms working in the steady state are similar to those off the steady state. The requirement of weakly separable production for the optimality of a zero tax on capital income has been shown by Stiglitz (1985) and Chari and Kehoe (1999). Homotheticity is not required in the steady state since consumption is constant by definition, that is, it is not possible to disproportionately tax some households by varying the tax rate in different periods. By definition, there are also no heterogeneous endowments in the steady state. Because of time-separable utility, heterogeneous endowments in earlier periods have no effect on the optimal capital income tax rate in the steady state. For a formal treatment, see appendix 2.5.1.1.

The above discussion has shown the importance of heterogeneity for optimal taxation in a dynamic setting. Optimal capital income taxes are zero if endowments are homogeneous, optimal present value labor income taxes are uniform if productivities are homogeneous. However, we cannot infer how the size and a possible correlation between these two sources of heterogeneity influence the determination of optimal tax rates. In addition, we cannot derive the impact of the planner's preferences, i.e. how do the determinants of optimal taxation change if the planner favors wealthier households? To address these issues we develop a more tractable model in the next section.

2.3 A Simplified Model

This section develops a two period model with two-dimensional heterogeneity. In order to focus on the interaction of different sources of heterogeneity, we use homothetic and identical preferences and weakly separable production.

Consider an economy with N households which are heterogeneous with respect to their labor productivities n^j and their non-negative endowments e^j, where $j = 1, .., N$. In each period, there is one consumption good (c_1^j, c_2^j) and one type of labor (l_1^j, l_2^j). Second period's utility is discounted with the factor ρ. Utility is log-linear and identical in every period and across households. That is, household j's utility in period 1 can be expressed as $u_1^j(c_1^j, l_1^j) = a\ln(1 - l_1^j) + (1 - a)\ln(c_1^j)$, where l_1^j is household j's labor supply in the first period, c_1^j is its consumption in the first period, and $a \in (0,1)$ determines the relative weights of leisure and consumption. Each household is endowed with one unit of time in every period. Total utility of household j is

$$U^j(c_1^j, l_1^j, c_2^j, l_2^j) = a\ln(1-l_1^j) + (1-a)\ln c_1^j + \rho\left[a\ln(1-l_2^j) + (1-a)\ln c_2^j\right] \quad . \tag{2.1}$$

In the first period, each household has to decide how much to work and how much to consume. It can save an amount k^j and will earn interest r on its savings in the second period. There is no depreciation of capital. The wage of household j is given by its productivity n^j and its labor income is $l_t^j n^j$ in each period $t = 1, 2$. Production is linear with first period's output given by $\sum_j n^j l_1^j$ and second period's by $\sum_j rk^j + n^j l_2^j$.

The government maximizes a social welfare function of the form $\sum_j \omega^j U^j$ through linear taxation of labor and capital income, where ω^j is the weight assigned to household j. Since the individual marginal utilities are decreasing in consumption and leisure, redistribution from wealthier households to poorer households increases welfare—unless the government favors richer households, implying that the welfare weights are positively correlated with individual utility. We assume that the government cannot observe endowments and productivities directly, it distinguishes the agents by their incomes only. The capital income tax is $(1-\beta_r)$ and the labor income tax $(1-\beta_w)$.[8] For the moment, we assume that labor taxes cannot be changed over time.

The households get lump-sum transfers α, which are identical in both periods and across households and may be either positive or negative. We assume that the government has a commitment technology. That is, once the households have made their labor/leisure decisions, the government cannot change the tax rates.[9] The households' budget constraint is:

$$c_1^j + (1-l_1^j)\beta_w n^j + \frac{c_2^j + (1-l_2^j)\beta_w n^j}{1 + r\beta_r} = inc^j \quad , \tag{2.2}$$

[8] If the government knows the amount of interest earnings and wage income of an individual, it would be reasonable to believe that it could infer the size of the initial endowment. We rule this out by assumption since heterogeneous endowments are meant to represent generic differences between individuals rather than purely monetary ones.

[9] This assumption is crucial for most work on optimal taxation. For an analysis of optimal taxation without commitment, see Klein and Rios-Rull (1999).

where $inc^j := \alpha + e^j + \beta_w n^j + \frac{\alpha+\beta_w n^j}{1+r\beta_r}$ denotes household j's present value potential income from its endowments in commodities (e^j) and work time ($\beta_w n^j + \frac{\beta_w n^j}{1+r\beta_r}$) plus transfers.

The government earns (pays) the same interest rate as the households on any budget surplus (deficit) in the first period.[10] Let B denote its budget, the government's budget constraint is

$$B = \sum_{j=1}^{N} \left\{ \left[(1-\beta_w) l_1^j n^j - \alpha\right](1+r) + (1-\beta_w) l_2^j n^j + (1-\beta_r) r k^j - \alpha \right\} \geq 0 \quad . \tag{2.3}$$

Given the tax rates and transfers, household j maximizes its utility subject to (2.2). Taking the households' optimal choices as given, the maximized utility of household j depends only on $\alpha, \beta_r, \beta_w, n^j$ and e^j and can be written as $V(\alpha, \beta_r, \beta_w, n^j, e^j)$. The indirect utility function $V(\cdot)$ is increasing in the household's productivity n^j and in its endowment e^j, since higher values of these variables lead to higher consumption and lower labor supply in both periods.

2.3.1 Optimal Taxation

This section derives properties of the optimal linear tax schedule with a special emphasis on capital income taxation. Given the households' choices, the planner chooses α, β_r, and β_w to maximize welfare. Letting W denote the corresponding Lagrangian, the planner's maximization problem can be written as

$$\max_{\alpha,\beta_r,\beta_w} W = \sum_{j=1}^{N} \omega^j V(\alpha, \beta_r, \beta_w, n^j, e^j) + \lambda B \quad , \tag{2.4}$$

where λ is the Lagrange multiplier for the planner's budget constraint. Since we want to show that a tax rate of zero is generally not optimal, we focus on the analysis of the planner's first order conditions evaluated at a tax rate of zero. That is, we calculate the marginal welfare effect of introducing labor or capital income taxes.

Let us first consider labor income taxes. The analysis of the necessary conditions for the maximization in (2.4) leads to the following Proposition:

[10] For ease of exposition, we assume that the government has no expenses besides redistribution. A fixed revenue requirement would not change the results since we allow lump-sum taxes.

PROPOSITION 4 *Given the utility and welfare functions specified above, and assuming that the weights ω^j are uncorrelated with productivities and endowments, the labor income tax $1-\beta_w$ is related to welfare as follows:*

i. A tax on labor income exceeding or equal to 100% ($\beta_w \leq 0$) is never optimal.

ii. Starting from a labor income tax rate of zero and assuming that $\beta_r > -\frac{1}{r}$, labor income taxes increase welfare if productivities are heterogeneous while endowments are not. The increase in welfare rises in the productivities' heterogeneity as measured by $\sum_j \frac{n^j-\bar{n}}{inc^j}$.[11]

iii. Starting from a labor income tax rate of zero and assuming that $\beta_r > -\frac{1}{r}$, labor income taxes decrease welfare if endowments are heterogeneous while productivities are not. The decrease in welfare rises in the endowments' heterogeneity as measured by $\sum_j \frac{e^j-\bar{e}}{inc^j}$.

Proof. If the labor income tax rate is 100% (*i.e.* $\beta_w = 0$) nobody works and the planner does not collect any revenue from labor taxation. Thus, a labor tax rate of 100% is never optimal, see also Sheshinski (1972).

For parts (*ii*) and (*iii*) consider the marginal welfare effect of labor income taxes. If there is no labor taxation ($\beta_w = 1$), the marginal effect is:[12]

$$\left.\frac{\partial W}{\partial \beta_w}\right|_{\beta_w=1} = \sum_{j=1}^{N} \omega^j(\rho+1)\left\{\frac{(1-a)(r\beta_r+2)\left[n^j-\bar{n}\right]-a(r\beta_r+1)\left[e^j-\bar{e}\right]}{inc^j}\right\} \quad , \quad (2.5)$$

where $inc^j = \alpha + e^j + \beta_w n^j + \frac{\alpha+\beta_w n^j}{1+r\beta_r}$. A negative value of this derivative implies that an increase in labor income taxes ($\beta_w \downarrow$) decreases welfare. If ω is uncorrelated with n and e, we can substitute $\bar{\omega}$ for ω^j in (2.5).

Assume $\beta_r > -\frac{1}{r}$. If endowments are homogeneous, $e^j = \bar{e}$ and the second term in the numerator is zero. The sign of (2.5) then is determined by the sign of $\sum_j (n^j-\bar{n})/(\alpha+\bar{e}+\beta_w n^j + \frac{\alpha+\beta_w n^j}{1+r\beta_r})$. If the numerator is positive ($n^j > \bar{n}$), the

[11] Throughout the remainder of the text, $\bar{x}$ refers to the arithmetic average of x, that is, $\bar{x} = \frac{1}{N}\sum_{j=1}^{N} x_j$.

[12] For the derivation calculate λ from $\frac{\partial W}{\partial \alpha} = 0$, plug it into $\frac{\partial W}{\partial \beta_w}$ and simplify.

denominator is larger than average. That is, positive values obtain a low weight. Negative values, on the other hand, obtain a high weight. As a consequence, the sum is negative. The higher the inequality in n, the lower the sum. This proves part (ii) of the Proposition.

If productivities are homogeneous, $n^j = \bar{n}$ and the first term in the numerator is zero. The sign of (2.5) then is determined by the sign of $-\sum_j (e^j - \bar{e})/(\alpha + e^j + \beta_w \bar{n} + \frac{\alpha+\beta_w \bar{n}}{1+r\beta_r})$. Using the same arguments as above, it follows that the sum is positive, rising in the inequality in e. This proves part (iii) of the Proposition. ■

Intuitively, if only productivities are heterogeneous more productive households work more than less productive ones and pay more taxes. Thus, at the margin, labor income taxation is redistributive and increases welfare. If only endowments are heterogeneous, the initial welfare effects of labor income taxes are negative, calling for wage subsidies and lump-sum taxes. Since wealthier households generally work less, they benefit less from subsidies while paying the same lump-sum tax.

Heterogeneity in e is measured by $\sum \frac{e^j - \bar{e}}{inc^j}$ which can be written as $Cov\left(e, \frac{1}{inc}\right) < 0$. If n and e are not correlated, increasing inequality with a mean preserving spread always implies a decrease in the covariance and, thus, an increase in the marginal welfare effects of taxation. A positive correlation between n and e further increases the marginal effects while a negative correlation decreases its value such that the marginal welfare effects of introducing labor taxes are eventually negative. This is the case if households with high endowments have a lower than average potential income ($inc^j < \overline{inc}$) because of their very low productivity.

If agents differ with respect to their productivities only, more productive agents have a higher income. This property is called "agent monotonicity" and implies that income taxes redistribute from highly productive agents to less productive agents since agents with high income also have a high productivity. Earlier work on optimal income taxes (Sheshinski, 1972, for example) has shown that, under this assumption, the optimal linear tax schedule consists of a positive transfer and a marginal income tax rate which is strictly positive and less than one. This result is consistent with part (ii) of Proposition 4, where agent monotonicity holds. In part (*iii*), however, agents with high endowments work less and, thus, generate lower labor income than agents with small endowments, making it optimal to subsidize labor income instead. If both endowments and productivities are heterogeneous and uncorrelated the optimality of either taxes or subsidies depends on the relative size of the heterogeneities and on the households' relative valuation of leisure and consumption, a.

If the welfare weights vary across households, their correlation with n and e is crucial for the determination of the optimal tax rate. A negative correlation, e.g. between the weights and endowments increases the marginal welfare effects of labor taxes, while a positive correlation decreases them and could even lead to opposite effects. Intuitively, if the government favors wealthier households, who generally work less than poorer households, labor subsidies are less desirable.

Now consider the tax rate on capital income, $1-\beta_r$. The following Proposition establishes the main relations between the capital income tax rate and welfare.

PROPOSITION 5 *Given the utility and welfare functions specified above, and assuming that welfare weights, productivities, and endowments are uncorrelated, the capital income tax* $1-\beta_r$ *is related to welfare as follows:*

i. *(a) A capital income tax exceeding 100% ($\beta_r < 0$) may be optimal if the inequality in endowments exceeds a lower bound for given weights w^j.*

(b) A confiscation of capital as well as interest ($\beta_r \leq -\frac{1}{r}$) is never optimal.

ii. Starting from a capital income tax rate of zero and assuming $\beta_w > 0$, capital income taxes increase welfare if either endowments are heterogeneous or productivities are heterogeneous and $\rho(1+r) > 1$, or both. The larger the heterogeneity as measured by $\sum_j \frac{e^j - \bar{e}}{inc^j}$ and $\sum_j \frac{n^j - \bar{n}}{inc^j}$, the larger the marginal welfare increase.

iii. Starting from a capital income tax rate of zero and assuming $\beta_w > 0$, capital income taxes decrease welfare if productivities are heterogeneous while endowments are not and $\rho(1+r) < 1$. The larger the heterogeneity as measured by $\sum_j \frac{n^j - \bar{n}}{inc^j}$, the larger the marginal welfare decrease.

Proof. If capital income taxes are 100% ($\beta_r = 0$), the marginal welfare impact of lowering capital income taxes ($\beta_r \uparrow$) is

$$\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=0} = \frac{r}{1+r+\rho}\sum_{j=1}^{N}\omega^j \cdot \left\{(1+\rho)\frac{\rho(e^j-\bar{e})-\beta_w(n^j-\bar{n})(1+r-\rho)}{inc^j}+r\rho\right\} \quad . \tag{2.6}$$

If ω, n and e are uncorrelated, we can replace ω^j by $\bar{\omega}$. The only negative term in (2.6) is $\sum_j \frac{e^j-\bar{e}}{inc^j}$. It is large in absolute terms if the inequality in endowments is very high. Thus, the derivative in (2.6) decreases in the endowments' inequality. Given the weights ω^j, $\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=0} = 0$ implicitly defines a lower bound for heterogeneity in endowments for which optimal capital income taxes may exceed 100%. This proves part (*ia*).

While the optimal capital income tax may exceed 100%, it is never optimal to tax away all savings. If $\beta_r \leq -\frac{1}{r}$, nobody saves and the planner does not collect any revenue from capital taxation. This proves part (*ib*).

For parts (*ii*) and (*iii*), consider the marginal welfare effects if there are no capital income taxes ($\beta_r = 1$):

$$\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=1} = r\sum_{j=1}^{N}\omega^j\frac{\rho(e^j-\bar{e})+\beta_w\left(\rho-\frac{1}{1+r}\right)[n^j-\bar{n}]}{(r+1)\cdot inc^j} \quad . \tag{2.7}$$

Again, a negative value of the derivative indicates that capital income taxes increase welfare.

The arguments here are similar to those given for Proposition 4. First, we can substitute $\bar{\omega}$ for ω^j. Second, the terms $\sum_j \frac{e^j-\bar{e}}{inc^j}$ and $\sum_j \frac{n^j-\bar{n}}{inc^j}$ are negative if n and e vary across agents and are not negatively correlated. Assume $\beta_w > 0$. If $\rho > \frac{1}{1+r}$, the partial derivative is negative and increasing in the heterogeneity of n and e, proving part (*ii*). If $\rho < \frac{1}{1+r}$, however, both sources of heterogeneity work in different directions. If only endowments are heterogeneous, (2.7) is negative and capital income taxes increase welfare. If only productivities are heterogeneous, (2.7) is positive and capital income taxes decrease welfare. This proves part (*iii*). ■

Let us spend a few more thoughts on (2.7). If productivity is the single source of heterogeneity, then $\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=1}$ is negative as long as $\rho > \frac{1}{1+r}$. The restriction on

ρ implies that the households' discount rate $\frac{1}{\rho} - 1$ is less than the interest rate r and is related to the households' optimal savings decision. If $\rho > \frac{1}{1+r}$, more productive households save more to shift more utility to the second period and, thus, pay a higher amount of capital income taxes than less productive households; redistribution occurs through capital income taxes and transfers. If $\rho < \frac{1}{1+r}$, more productive households save less and, thus, gain less from capital income subsidies than less productive households; redistribution occurs through capital income subsidies. If only endowments are heterogeneous the argument is similar: households with higher endowments save more and, thus, pay more capital income taxes than households with lower endowments. If both endowments and productivities are heterogeneous and if $\rho > \frac{1}{1+r}$, the effects once again depend on the correlation between n and e. The marginal welfare impacts are stronger for positive correlation, weaker or even reversed for negative correlation.

To summarize, we found that the influence of heterogeneity on the optimal tax rates depends strongly on the source of the heterogeneity and on possible correlations between the different sources. While, at the margin, labor taxes are welfare enhancing if productivities are heterogeneous, they can reduce welfare if endowments are heterogeneous. Capital income taxes increase welfare if endowments are heterogeneous while the effect of heterogeneous productivities depends on the sign of $\rho - \frac{1}{1+r}$. If welfare weights vary across households, the marginal welfare effects crucially depend on the weights' correlation with the households' endowments and productivities. The results confirm the intuition: if the government favors well-to-do households, marginal welfare effects of taxation are lower; if it favors poorer households, they are higher.

2.3.2 Optimal Taxation with Time-dependent Labor Taxes

Up to now, the planner was restricted to tax labor income in both periods with the same tax rate $1 - \beta_w$. The present section modifies the above analysis to allow time-varying labor income taxation. $(1 - \beta_{w1})$ is the first period's labor income tax and $(1 - \beta_{w2})$ is the second period's labor income tax. As before, we assume that the government credibly commits to second period's labor and capital income taxes before households make their labor/leisure choices. What are the effects on the relation between capital income taxes and welfare?

PROPOSITION 6 *Consider the environment as described above, with labor income taxes free to vary between both periods. If labor taxes are at their optimal values, then, starting from a capital income tax rate of zero:*

i. An increase in the capital income tax rate has no first order effects if endowments are homogeneous.

ii. An increase in the capital income tax rate increases welfare if endowments are heterogeneous and the correlation between n and e is not too negative, given the weights ω^j.

iii. An increase in the capital income tax rate decreases welfare if endowments and productivities are heterogeneous and their correlation is sufficiently negative, given the weights ω^j.

Proof. If there are no capital income taxes and if labor taxes are at their optimal values, the marginal welfare impact of capital income taxation is[13]

$$\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=1} = \frac{r\rho}{1+r} \cdot \sum_{j=1}^{N} \left\{ \frac{\omega^j (e^j - \bar{e})}{inc^j} \right\} \quad . \tag{2.8}$$

It follows directly that $\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=1} = 0$ if endowments are homogeneous, proving part (i). The derivative is negative if e is heterogeneous and not correlated with n and ω. The derivative is positive only if there is a sufficiently negative correlation since negative values of the sum ($e^j < \bar{e}$) are associated with very high values of n^j leading to lower than average weights ($inc^j < \overline{inc}$). The critical level is implicitly given by $\sum \frac{\omega^j (e^j - \bar{e})}{inc^j} = 0$. This proves parts ($ii$) and ($iii$) of the Proposition. ■

Proposition 6 indicates that the optimality of a positive tax rate on capital income is solely driven by heterogeneity in endowments. This result is analogous

[13] For the derivation of this expression see appendix 2.5.2.

to Proposition 1, where we have shown that optimal capital income taxes can be zero if production is weakly separable, preferences are identical and homothetic, and there are no heterogeneous endowments. Equation (2.8), however, also shows how correlations between productivities, endowments and welfare weights change the marginal welfare effects of capital income taxation. At the margin, capital income taxes increase welfare if n and e are not correlated. If they are positively correlated, the marginal welfare effect of capital income taxation increases since households with higher interest earnings tend to be more productive as well and thus have a higher income than others. However, if they are negatively correlated, the size of the marginal welfare improvement decreases since households with higher endowments tend to be less productive. For very high levels of negative correlation—where households with higher endowments tend to have lower overall utility than others because of their low productivity—it is desirable to pay interest subsidies instead of imposing taxes. These results are compatible with the findings of Domeij and Heathcote (2000), who examine the quantitative effects of eliminating capital income taxes. They find that if households are heterogeneous, a vast majority prefers the original tax system. If the population is homogeneous, however, a tax cut would be preferred.

How do welfare weights influence the marginal effects? A negative correlation between weights and endowments or productivities—implying that the government favors poorer households—strengthens the marginal welfare improvement of capital income taxation. A positive correlation, on the other hand, lowers the positive impact of capital income taxation. When comparing these results with the previous section where labor taxes were restricted to be constant over time we find that the expressions determining the marginal welfare impact of capital income taxes are very similar (equations 2.7 and 2.8). Optimal capital income taxes are largely determined by the correlation between endowments and productivities in both cases. If the planner is restricted to identical labor taxes in both periods, capital income taxes can be employed to substitute for the loss of flexibility. That is, even if endowments are homogeneous, the planner might find it optimal to tax interest income. If he can vary labor income taxes over time, it is not optimal for him to tax capital income.

2.4 Conclusion

The analysis in the previous sections has discussed the influence of heterogeneity among households on optimal tax rates. While steady state analyses, such as Judd (1985), have found that the optimal capital income tax rate in the steady state is zero, we have shown that off the steady state this is generally not the case.

For an economy with arbitrarily many heterogeneous households, we have shown that if households are heterogeneous with respect to productivities and endowments, zero capital income taxes generally are not optimal. Emphasizing the analogy between commodity taxes and capital income taxes, we have found that zero capital income taxation can be optimal only if endowments are homogeneous, if production is weakly separable between labor and capital, and if utility functions are homothetic and identical across agents. In a simplified model we have shown that the extent of the inequality and the joint distribution of its different components (productivities and endowments in our model) are crucial for the size of the marginal welfare effects of taxation. A positive correlation between endowments and productivities increases the marginal welfare effects of capital income taxation, while a negative correlation decreases the effects. Correlation of the households' characteristics, thus, plays an important role in determining the optimal tax policy.

2.5 Appendix

2.5.1 Derivation of Propositions 1 to 3

The following model is an extension of Bassetto (1999).[14] There are N households, preferences are separable between consumption and leisure. Household j's utility is given by $U^j(G^j(c^j), H^j(1-l^j))$, where $G^j(\cdot)$ is its subutility from consumption, and $H^j(\cdot)$ is its subutility from leisure.[15] $c^j = (c_1^j, c_2^j, ..)$ is the vector of its consumption and $l^j = (l_1^j, l_2^j, ..)$ is the vector of its time spent working, where subscripts refer to the time period. The endowment of time is one for each household in every period. The intertemporal technology constraint is given by $F\left(\sum_{j=1}^N c^j + g, l^1, .., l^N\right) \leq 0$, where g is government consumption and $F(\cdot)$ is assumed to be twice continuously differentiable, increasing in the first argument and decreasing in labor. In the following, we use the "primal approach" or "Ramsey approach" to determine properties of optimal tax rates.[16]

The firm produces consumption goods c_t, sells them at a price q_t and pays wages w_t^j. Wages are per unit of time and differ across households. The firm solves

$$\max_{c_t, l_t^j} \sum_t \left\{ q_t c_t - \sum_j w_t^j l_t^j \right\} \quad \text{s.t.} \quad F\left(\sum_{j=1}^N c^j + g, l^1, .., l^N\right) \leq 0 \quad . \tag{2.9}$$

From the firm's first order conditions, $q_t = -F_{c_t}/F_{l_1^1}$ and $w_t^j = F_{l_t^j}/F_{l_1^1}$, where w_1^1 is normalized to 1. Let $\tau_{wt} w_t^j$ denote household j's after tax wages, p_t the consumer price for goods, and e_t^j household j's endowment in period t, and α the government transfers (which are distributed in the initial period only), then household j maximizes

$$\max_{c_t^j, l_t^j} U^j(G^j(c^j), H^j(1-l^j)) \ \text{ s.t. } \ \sum_t p_t(c_t^j - e_t^j) - \sum_t \tau_{wt} w_t^j l_t^j - \alpha = 0 \quad . \tag{2.10}$$

From the first order conditions, it follows that $\frac{G_t^j}{G_1^j} = \frac{p_t}{p_1}$, $\frac{H_t^j}{H_1^j}\frac{F_{l_t^1}}{F_{l_t^j}} = \frac{\tau_{wt}}{\tau_{w1}}$, and $\frac{\tau_{wt}}{p_t} = \frac{U_2^j H_t^j}{U_1^j G_t^j}\frac{F_{l_1^1}}{F_{l_t^j}}$, where functions with subscripts (U_k, H_k, G_k) denote partial derivatives.

[14] Bassetto (1999) considers two agents, of whom only one works. We extend his model to an arbitrary number of heterogeneous agents who are all working.

[15] In what follows, t, k refer to different periods, while j refers to households.

[16] For a comprehensive overview see Chari and Kehoe (1999).

Normalizing p_1 to 1 we can write the implementability constraint as

$$U_1^j \sum_t G_t^j (c_t^j - e_t^j) - U_2^j \sum_t H_t^j l_t^j - U_1^j G_1^j \alpha \leq 0 \quad . \tag{2.11}$$

Since all agents face the same consumption prices, marginal rates of substitution must be equal for all households in competitive equilibrium. The analogy holds for wages, given heterogeneous productivities. That is,

$$G_t^1 G_1^j = G_1^1 G_t^j \quad \text{and} \quad \frac{H_t^1 F_{l_1^1}}{H_t^j F_{l_1^j}} = \frac{H_1^1 F_{l_t^1}}{H_1^j F_{l_t^j}} \quad \forall\, t, j \quad . \tag{2.12}$$

After using the households' necessary conditions to manipulate the implementability constraint (2.11), the planner's maximization problem can be written as

$$\begin{aligned}
\max_{c_t^j, l_t^j, \alpha} \quad & \sum_j \omega^j U^j \left(G^j(c^j), H^j(1-l^j)\right) - \mu F\left(\sum_j c^j + g, l^1, .., l^N\right) \\
& + \sum_j \lambda^j \left[U_1^1 \sum_k G_k^1 (c_k^j - e_k^j) - U_2^1 \sum_k \frac{F_{l_k^j}}{F_{l_k^1}} H_k^1 l_k^j - U_1^1 G_1^1 \alpha\right] \\
& + \sum_{j>1} \sum_{k>1} \left\{\nu_k^j \left[G_1^j G_k^1 - G_k^j G_1^1\right] + \eta_k^j \left[\frac{H_k^1 F_{l_1^1}}{H_k^j F_{l_1^j}} - \frac{H_1^1 F_{l_k^1}}{H_1^j F_{l_k^j}}\right]\right\} \quad ,
\end{aligned}$$

where $\lambda^j, \nu_k^j, \eta_k^j$, and μ are Lagrange multipliers.

2.5.1.1 Proof of Propositions 1 and 3

This section derives the properties of optimal commodity tax rates over time and closely follows Bassetto (1999). Commodity taxes are determined by $\frac{p_t}{q_t} = \frac{G_t^j}{F_{c_t}} \frac{F_{c_1}}{G_1^j} \frac{1}{q_1}$ $\forall j, t$, and commodity taxation is uniform if $\frac{G_t^j}{F_{c_t}}$ is independent of t. The following paragraphs examine under which conditions uniform commodity taxation is compatible with the planner's necessary conditions.

Using the fact that the first order condition w.r.t. α implies $\sum_j \lambda^j = 0$, the derivative of the welfare function with respect to c_t^1 can be written as

$$c_t^1 \quad : \quad \frac{G_t^1}{F_{c_t}} \left\{\left(\omega^1 + \lambda^1\right) U_1^1 + U_{11}^1 \sum_j \sum_k \lambda^j G_k^1 \left(c_k^j - e_k^j\right)\right\} \tag{2.13}$$

$$= \mu - \frac{1}{F_{c_t}} \left[U_1^1 \sum_j \sum_k \lambda^j G_{kt}^1 \left(c_k^j - e_k^j \right) - U_2^1 \sum_{j>1} \sum_k \lambda^j H_k^1 l_k^j \Lambda_{k,c_t}^{j,1} \right.$$
$$\left. + \sum_{j>1} \sum_{k>1} \left\{ \nu_k^j \left[G_1^j G_{kt}^1 - G_k^j G_{1t}^1 \right] + \eta_k^j \left[\frac{H_k^1}{H_k^j} \Lambda_{1,c_t}^{1,j} - \frac{H_1^1}{H_1^j} \Lambda_{k,c_t}^{1,j} \right] \right\} \right] \quad ,$$

where $\Lambda_{k,c_t}^{1,j} = \partial \left(\frac{F_{l_k^1}}{F_k^j} \right) / \partial c_t$. The r.h.s is independent of t if the following terms are equal to zero:

- $\sum_j \sum_k \lambda^j G_{tk}^1 c_k^j$. If G^1 is linearly homogeneous, G_t^1 is homogeneous of degree 0 and $\sum_k G_{tk}^1 c_k^1 = 0$. If, in addition, the functions G^1 and G^j are identical, c^1 and c^j are proportional if preferences are homothetic. Thus, the above sum is zero. An alternative possibility would be identical consumption for all types of agents.

- $\sum_j \sum_k \lambda^j G_{tk}^1 e_k^j$. Since $\sum_j \lambda^j = 0$, this expression is zero if a) endowments are homogeneous or b) endowments are proportional to actual consumption and the previous term is zero.

- $\Lambda_{k,c_t}^{1,j}$ and $\Lambda_{k,c_t}^{j,1}$. These terms equal zero if either productivities are identical for all agents or if $F(\cdot)$ is weakly separable between $(l_k^1, .., l_k^N)$ and c_t.[17]

- $\sum_{j>1} \sum_{k>1} \nu_k^j (G_1^j G_{tk}^1 - G_k^j G_{1t}^1)$. ν_k^j is the multiplier on the equality of marginal rates of substitution. We can show that the constraint is not binding if the previous conditions hold, implying that $\nu_k^j = 0 \ \forall k, j$.

Assume that the above conditions hold and that $\nu_k^j = 0 \ \forall k, j$. The first order conditions with respect to consumption then can be expressed as

$$c_t^1 \ : \ \frac{G_t^1}{F_{c_t}} \left\{ \left(\omega^1 + \lambda^1 \right) U_1^1 + U_{11}^1 \sum_j \sum_k \lambda^j G_k^1 (c_k^j - e_k^j) \right\} = \mu \tag{2.14}$$

$$c_t^j \ : \ \frac{G_t^j}{F_{c_t}} \left\{ \omega^j U_1^j \right\} + \frac{G_t^1}{F_{c_t}} \left\{ \lambda^j U_1^1 \right\} = \mu \qquad \forall j > 1 \ . \tag{2.15}$$

As a consequence of (2.14), G_t^1 / F_{c_t} does not depend on t. From (2.15) we then see that, for any j, G_t^j / F_{c_t} does not depend on t either. Thus, G_t^1 / G_t^j is constant $\forall \ j$. Consequently, the constraint on the equality of marginal rates of substitution (2.12) is not binding and $\nu_k^j = 0 \ \forall k, j$.

[17] This is equivalent to production being weakly separable between labor and capital.

Weakly separable utility If utility is denoted by $U^j(c^j, 1-l^j)$, weak separability implies $U^1_{c_t,l_k} = \frac{U^1_{c_1,l_k}}{U^1_{c_1}} U^1_{c_t}$. Using this relation and going through the same steps as above, one finds that the only difference to the earlier analysis with strongly separable preferences is that preferences need to be homothetic and identical for all elements of the utility function.

Time-separable utility If utility is time separable, $U_{11} G_t G_k + U_1 G_{tk} = 0$ or $G_{tk} = -G_k G_t \frac{U_{11}}{U_1}$ $\forall\ k \neq t$. The analysis follows the same steps as in section 2.5.1.1, the only difference concerning the following term:

- $\sum_j \lambda^j \sum_k G^1_{kt} e^j_k$. If utility is time separable, $\sum_k G^1_{kt} = -\sum_k \left\{ G^1_k G^1_t \frac{U^1_{11}}{U^1_1} \right\} + (G^1_t)^2 \frac{U^1_{11}}{U^1_1} + G^1_{tt}$ and

$$\sum_j \lambda^j \sum_k G^1_{kt} e^j_k = -G^1_t \left[\frac{U^1_{11}}{U^1_1} \sum_j \lambda^j \sum_k G^1_k e^j_k \right] + \left[(G^1_t)^2 \frac{U^1_{11}}{U^1_1} + G^1_{tt} \right] \sum_j \lambda^j e^j_t \quad .$$

The first term on the r.h.s is a multiple of G^1_t and moves to the l.h.s. of (2.13). The second term is zero if endowments in period t are homogeneous since $\sum_j \lambda^j = 0$.

In the steady state As in a steady state, assume that after period $\kappa < t$ consumption and labor are constant and there are no endowments. Since utility is time-separable, G_t and G_{tt} are constant. Again, the analysis follows the same steps as above with slight changes for the following two terms.

- $\sum_j \lambda^j \sum_k G^1_{kt} e^j_k$. For $t > \kappa$ there are no endowments and the argument is the same as above.

- $\sum_j \lambda^j \sum_k G^1_{kt} c^j_k$. If utility is time separable it follows that

$$\sum_j \lambda^j \sum_k G^1_{kt} c^j_k = -G^1_t \left[\frac{U^1_{11}}{U^1_1} \sum_j \lambda^j \sum_k G^1_k c^j_k \right] + \left[(G^1_t)^2 \frac{U^1_{11}}{U^1_1} + G^1_{tt} \right] \sum_j \lambda^j c^j_t \quad .$$

Again, the first term moves to the l.h.s. of (2.13). For $t > \kappa$ the second term is time invariant since consumption is constant.

2.5.1.2 Proof of Proposition 2

This section derives properties of optimal labor income taxes over time. Labor income taxes are determined by $\tau_{wt} = \frac{U_2^j}{U_1^j}\frac{p_1}{w_1^1}\frac{F_{l_1^1}}{G_1^j} \cdot \frac{H_t^j}{F_{l_t^j}}$. They are uniform if $\frac{H_t^j}{F_{l_t^j}}$ is independent of t. The first derivative of welfare with respect to l_t^1 can be written as

$$\begin{aligned} l_t^1 \; : \quad & -\frac{H_t^1}{F_{l_t^1}}\left\{\omega^1 U_2^1 + \lambda^1 U_2^1 - U_{22}^1 \sum_j \lambda^j \sum_k H_k^1 l_k^j \Lambda_k^{j,1}\right\} \\ & = \mu + \frac{1}{F_{l_t^1}}\left\{U_2^1\left[\sum_j \lambda^j \sum_k H_{kt}^1 l_k^j \Lambda_k^{j,1} + \sum_{j>1} \lambda^j \sum_k H_k^1 l_k^j \Lambda_{k,l_t^1}^{j,1}\right]\right. \\ & \left. - \sum_{j>1}\sum_{k>1} \eta_k^j \left[\frac{H_{kt}^1}{H_k^j}\Lambda_1^{1,j} + \frac{H_k^1}{H_k^j}\Lambda_{1,l_t^1}^{1,j} - \frac{H_1^1}{H_1^j}\Lambda_{k,l_t^1}^{1,j} - \frac{H_{1t}^1}{H_1^j}\Lambda_k^{1,j}\right]\right\} \quad , \end{aligned} \tag{2.16}$$

where $\Lambda_{k,l_t^1}^{j,1} = \partial\left(\frac{F_{l_k^j}}{F_k^1}\right)/\partial l_t^1$. The independence of $\frac{H_t^1}{F_{l_t^1}}$ from t is violated by the following terms.

- $\sum_j \sum_k \lambda^j H_{kt}^1 l_k^j \Lambda_k^{j,1}$. If the difference in productivities is identical for all goods, $\Lambda_k^j = \Lambda^j \; \forall j, k$. If H is linearly homogeneous and identical across agents, this term can be modified to $-\sum_j \lambda^j \Lambda^j \sum_k H_{kt}^1$ since $\sum_k H_{kt}^1(1 - l_t^j) = 0$. From the maximization with respect to α we know that $\sum_j \lambda^j = 0$. Thus, if Λ^j is independent of j, that is, if productivities are homogeneous, this term equals zero.
- $\sum_{j>1}\sum_k \lambda^j H_k^1 l_k^j \Lambda_{k,l_t^1}^{j,1}$. If productivities are homogeneous, $\Lambda_{k,l_t^1}^{j,1}$ is zero.
- $\sum_{j>1}\sum_{k>1} \eta_k^j \left[\frac{H_{kt}^1}{H_k^j}\Lambda_1^{1,j} + \frac{H_k^1}{H_k^j}\Lambda_{1,l_t^1}^{1,j} - \frac{H_1^1}{H_1^j}\Lambda_{k,l_t^1}^{1,j} - \frac{H_{1t}^1}{H_1^j}\Lambda_k^{1,j}\right]$. This term is zero if $\eta_k^j = 0 \; \forall \; k, j$. In analogy to the proof of Proposition 1 above, one can show that if the above conditions hold the respective constraint is not binding, that is, $\eta_k^j = 0 \; \forall \; k, j$.

Weakly separable utility If utility is denoted by $U^j(c^j, 1 - l^j)$, weak separability implies $U_{l_t,c_k} = \frac{U_{l_1,c_k}}{U_{l_1}} U_{l_t}$. Using this relation and going through the same steps as above, one finds that the only difference to the earlier analysis with strongly separable preferences is that preferences need to be homothetic and identical for all elements of the utility function.

2.5.2 Derivation of the Partial Derivative with Respect to β_r

The partial derivatives of the planner's welfare function are modified the following way. Firstly, calculate λ from $\frac{\partial W}{\partial \alpha} = 0$. Secondly, manipulate the other partial derivatives and substitute for λ to get

$$\frac{\partial W}{\partial \beta_r}\frac{1+r\beta_r}{r} + \frac{\partial W}{\partial \beta_{w2}}\frac{\beta_{w2}}{\rho+1} - \frac{\partial W}{\partial \beta_{w1}}\frac{\rho\beta_{w1}}{\rho+1} =$$

$$\rho \cdot \sum_{j=1}^{N} \omega^j \left\{ \frac{e^j - \bar{e}}{inc^j} + \frac{r(1-\beta_r)}{X(\rho+1)} \left(\frac{r\beta_r+2}{r\beta_r+1}\right) \left[(1-a)\frac{\overline{inc}}{inc^j} + \frac{2a}{inc^j}\left[\alpha + \bar{e} + \frac{\alpha}{r\beta_r+1}\right]\right]\right\} \quad ,$$

$$X = (2+r) + \frac{a}{\rho+1}\left[\frac{(1-\beta_{w1})}{\beta_{w1}}\frac{1+r}{r\beta_r+1} + \frac{1-\beta_{w2}}{\beta_{w2}}\rho\right](r\beta_r+2) - r(1-\beta_r)\frac{\rho - \frac{1}{1+r\beta_r}}{\rho+1} \quad .$$

If β_{w1} and β_{w2} are at their optimal values, the respective partial derivatives are zero and one can easily calculate $\frac{\partial W}{\partial \beta_r}$ from the above equation.

Part II

Improving Access to Financial Institutions through Microfinance

Chapter 3

Investigating Microfinance: Caja Los Andes, Bolivia

3.1 Introduction

During the last 20 years, microfinance—a new concept for the provision of credits to otherwise credit constrained poor households—has spread to most parts of the world. Compared to previous attempts to provide credit to the poor, the novelty of microfinance consists in a) the use of new incentive mechanisms such as group loans or the choice of collateral based on the borrower's subjective valuation, and b) the—more or less successful—attempt to cover costs through high interest rates which at the same time make these loans unattractive to better off borrowers. A recent overview is provided in Morduch (2000).

Although microfinance is highly popular among donors of development aid, the impact of these loans is not very well documented. Endogeneity of program placements and the absence of data on rejected loan applicants make it hard to find good control groups for a rigorous econometric analysis.[1] Acknowledging these problems, the present study is limited to the analysis of the clients from one microlender: Caja Los Andes in Bolivia. The foremost purpose of the following sections is to provide a description of the data set with a focus on the development of the clients over time.

[1]See Pitt and Khandker (1998) for a discussion of these issues.

	Branch of observation							Gross Loan
Year	La Paz	Sucre	Cochab.	Tarija	Trinidad	St. Cruz	Total	Portfolio
12/1992	1,518						1,518	0.7
12/1993	4,128						4,128	1.4
12/1994	9,846	309					10,155	2.9
12/1995	13,704	1,673	1,541	730			17,648	6.0
12/1996	18,771	2,465	3,615	1,785			26,636	11.7
12/1997	22,318	2,909	4,082	2,341	538		32,182	20.3
12/1998	24,008	3,319	4,235	3,393	1,432		36,387	28.4
12/1999	25,237	4,091	5,667	4,181	2,003	471	41,650	35.6
06/2000	23,207	3,804	6,546	4,105	1,749	911	40,322	40.8

Source: own calculations and IPC GmbH (2000).

Table 3.1: Number of active clients by branch and gross loan portfolio over time. A client is active if he or she has a loan outstanding at some time during the respective year. The gross loan portfolio is in million $US.

The data analyzed is from various parts of Bolivia. The microfinance market in Bolivia is one of the most developed and most competitive microfinance markets in existence today, its characteristics thus might serve as an indicator for future trends in other markets (Von Stauffenberg 2001). The data set from Caja Los Andes is provided through the Interdisziplinäre Project Consult (IPC) GmbH in Frankfurt, who has supported Caja Los Andes and its predecessor Pro-Credito since 1992. The data consists of time series on individual borrowers including details on the loans taken, personal information, and information about the clients' businesses.

The remainder of this paper is organized as follows. Section 3.2 provides a brief overview of Caja Los Andes and microfinance in Bolivia. A descriptive analysis of the data set then is given in section 3.3, while section 3.4 provides a preliminary examination of a few hypotheses.

3.2 General Information about Bolivia

3.2.1 Caja Los Andes

Caja Los Andes FFP S.A. is a registered savings and loan company with its main branch in La Paz, Bolivia.[2] Its operations are dedicated to the provision of sustainable financial services to the economically disfavored. In July 2000 (when the data was collected) it offered credits to small and micro enterprises in rural and urban areas and also to the general public, the latter being secured by gold pawning. It also offered savings accounts and fixed deposits. Both loans and deposits are either in Bolivianos or in $US, acknowledging the widespread dollarization of the Bolivian economy.

In December 1999, Caja Los Andes was serving 36,815 clients. 39,335 loans were outstanding amounting to $US 35.9 Mio. 54% of these loans were made to women. The high concentration of micro-enterprises in the commerce sector is mirrored in the distribution of the outstanding loans, 44% of which went to commerce, 21% to production, 12% to the service industry, and 12% to agriculture related businesses. These numbers represent the outstanding loans in all branches, among which there are considerable differences. Agriculture, for example, plays a dominant role in rural branches, amounting to 70% of all outstanding loans there. For more details see Caja Los Andes (1999) and tables 3.23 and 3.24 in the appendix for the main branch in La Paz.

The origins of Caja Los Andes go back to a non-profit organization, Pro-Credito, which was founded in 1992 to provide credits to poor households in La Paz, Bolivia. During the first five years it received technical support from IPC GmbH, Germany. The funds provided under this contract became the most important source of equity for Caja Los Andes. Soon after the beginning of its operations, branches in Sucre (1994), Cochabamba and Tarija (1995) were opened. In 1995 Pro-Credito transformed to a registered private savings and loan company. The formal registration

[2]FFP stands for Fondo Financiero Privado which is a registered savings and loan company subject to a US$ 1 Mio. minimum equity requirement and certain restrictions on assets and liabilities. The legal category FFP has been created as an institutional form for small banks (Rhyne 2001).

Institution	Amount in $US
International/Foreign Institutions	
BID/FOMIN, CAF, PROFUND, HIVOS-TRIODOS, FMO, IFC	10,373,547
Subordinate obligations	
PRO-CREDITO, PROFUND	1,596,644
National second tier institutions	
FUNDA-PRO, NAFIBO, FCD, Prefectura del Beni	12,362,999
National financial institutions	1,741,392
Deposits from the public and from other institutions	8,675,556
Total	34,750,137

Source: Caja Los Andes (1999)

Table 3.2: Sources of capital, 12-31-1999.

made it possible to take deposits and also facilitated growth through the acquisition of long-term financing from international financial markets. The number of clients and the amount of loans disbursed increased considerably and more branches were opened, see table 3.1. The number of new clients per year has increased from 1,518 in 1992 to 11,831 in 1999, see table 3.22 in the appendix. Since many clients take repeat loans, the number of active clients is considerably higher, it rose from 1,518 in 1992 to 41,650 in 1999. While the main branch in La Paz still accounts for the largest fraction of this increase, the younger branches together served about 45% of the new clients and 40% of all active clients in 1999. In the first half of 2000, Caja Los Andes served 40,332 clients, 17% of whom were new customers. This growth is mirrored in the number of loans disbursed, displayed in table 3.21, and in the rise of the gross loan portfolio from $US 0.7 Mio. in 1992 to $US 40.8 Mio. in June 2000 as shown in table 3.1.

The fast growth rates made it necessary to tap new sources of capital. During recent years the liabilities to national second tier institutions and international institutions have particularly grown (from $US 2.1 Mio. in 1997 to $US 12.4 Mio. in 1999 and from $US 3.1 Mio. to $US 10.4 Mio, respectively), see Caja Los Andes (1999). For the current composition of Caja Los Andes' liabilities see table 3.2. The formal registration allowed to offer savings accounts and fixed deposits as well. From only 647 accounts in 1996 the number increased to 11,550 in December 1999.

year	staff	out. loan contracts / staff	gross loan portfolio / staff	admin. costs	interest costs	arrears >30 days	equity	roe after taxes
1995	94	170	64,000	27%	9.5%	0.5%	1.237	3%
1996	110	263	108,000	20%	9.0%	0.5%	1.824	18%
1997	143	207	143,000	14%	9.3%	0.8%	2.809	38%
1998	213	164	134,000	13%	9.1%	1.8%	3.686	27%
1999	270	146	133,000	13%	8.5%	3.8%	4.139	14%

Source: IPC GmbH (2000).

Table 3.3: Business indicators for Caja Los Andes. Costs are relative to average gross portfolio. roe = return on equity.

While most of the savings accounts contain very small sums only with an average of $US 99 in 1999, the average fixed deposit was $US 50,689.

Throughout its time of operation, Caja Los Andes has improved its efficiency and generated considerable profits. An overview of these indicators for the years 1995 to 1999 is provided in table 3.3. The administrative costs as percentage of the average gross portfolio decreased from 27% to 13% and the average outstanding loan portfolio per staff member increased from $US 64,000 to $US 133,000. The decreasing costs are partly passed on to the clients in the form of lower interest rates, contributing to the decline of income per average gross portfolio from 40% in 1995 to 29% in 1999. While most indicators show considerable improvement during recent years, arrears rates have increased from 0.5% in December 1995 to 3.8% in December 1999 and to 7.3% in June 2000, see IPC GmbH (2000). Arrears rates, however, have increased considerably in all parts of Bolivia's banking system due to a recession beginning in late 1998, see also section 3.2.2.2 and Von Stauffenberg (2001).

While the bank initially gave loans to micro-enterprises only (i.e. very small enterprises), the target group has broadened in recent years. Caja Los Andes has opened a new small enterprise division which gives substantially larger loans. Between January 1999 and June 2000 these loans made up 3.9% of all loans and about 14% of the amount disbursed.[3] Table 3.4 displays the median and the mean of the

[3] A loan belongs to the small enterprise group when the business has assets of at least $US

	Branch of observation						
Year	cbb	lpb	scz	sre	tdd	tja	Total
1992		367					367
		624					624
1993		285					285
		545					545
1994		205		161			204
		427		294			424
1995	232	238		193		195	232
	354	510		346		290	472
1996	332	271		267		271	272
	522	587		441		404	551
1997	430	347		344	419	339	356
	738	803		616	728	519	755
1998	483	470		440	404	440	469
	887	995		802	799	742	929
1999	465	525	586	437	464	520	511
	935	1,131	829	845	882	1,006	1,037
June 2000	495	602	692	428	432	595	560
	913	1,293	946	885	837	1,173	1,144

Table 3.4: Median and mean of the amount of disbursed loans by year and branch. Values are in 1992 \$US. cbb=Cochabamba, lpb = La Paz, scz = Santa Cruz, sre = Sucre, tdd = Trinidad, tja = Tarija.

loan amounts handed out.[4] The data show an increase in both values. For further characteristics of disbursed loans see section 3.3.2 and tables 3.23 and 3.24 in the appendix for July 2000 data of the La Paz branch.

20,000 and a monthly sales revenue of at least \$US 6,000. In July 2000, small enterprise loans were disbursed in the non-agricultural sectors in La Paz, Cochabamba, and Tarija only.

[4]The original data is partly in Bolivianos, partly in \$US. I have converted Bolivianos to \$US using daily exchange rates and then deflated it to 1992 using the US GNP deflator provided by the Federal Reserve Bank.

3.2.2 Economic Environment

This section briefly describes important characteristics of the Bolivian economy. Section 3.2.2.1 presents basic economic and social data of the 1990's and section 3.2.2.2 describes the microfinance market in Bolivia with a special emphasis on recent developments.

3.2.2.1 Economic and Social Information about Bolivia

Bolivia is one of the poorest countries in South America. After economic reforms and continuing market liberalization in the 1980's, the early nineties were characterized by relatively high GDP growth with declining unemployment, see figures 3.1 and 3.2. Beginning in late 1998 and continuing throughout 1999 and 2000, however, an economic crisis emerged leading to severe job cuts. In part, this crisis was caused by the devaluation of the Brazilian Real making Bolivian exports less competitive, by an increase of trade restrictions by the main Bolivian trading partners, Chile and Argentina, and by the large scale eradication of coca plantations. In 1999 Bolivian GNP per capita was $US 1,000 (Worldbank 2000a), corresponding to $US 3000 after purchasing power adjustment (CIA 1999). Wages are low; in 2000 police officers and teachers went on strike for wage increases to a monthly level of $US 100. While inflation has been moderate in the 1990's (generally below 10%, 3.3% in 1999 (Worldbank 2000a)), there is continuing devaluation against the $US, see figure 3.3. During the time frame covered by our data set the exchange rate nearly doubled from 3.75 Bolivianos per $US on January 31st 1992 to 6.17 on July 31st 2000.

Bolivia is landlocked and covers a wide variety of different climates, the altitude ranging from 90m to 6,542m. The infrastructure is poorly developed, the high costs due to the difficult terrain slow down improvements. Inequality is high with a Gini coefficient for income of 0.52 (1993, Inter-American Development Bank (1999)). There is widespread illiteracy with an overall rate of 16.4% and a rate of 22% among women (1997, Inter-American Development Bank (1999)).

The Bolivian economy is characterized by a large micro-enterprise sector. In study conducted by the Inter-American Development Bank (IADB), Orlando and

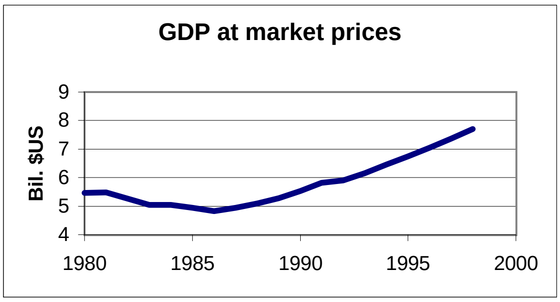

Figure 3.1: Bolivian Gross Domestic Product at market prices in $US. Source: IADB.

Pollack (2000) report that in the mid 1990's the Bolivian micro-enterprise sector represented 57% of total employment (compared to 54% in all of Latin America).[5] Most micro-enterprise workers are self-employed. Between 1990 and 1995, 84% of all new jobs in Latin America were created in the micro-enterprise sector. This sector is characterized by low wages, low human capital, and relatively large poverty. Women's earnings in the micro-enterprise sector are considerably lower than men's. In Bolivia, women obtain roughly 53% of men's earnings. While the self-employed microentrepreneurs earn more than workers in the formal sector, employed workers in micro-enterprises gain very low wages and a large proportion lives in poverty. Among the micro-enterprises the industry sector has the most poor earners, the commerce sector the fewest. The sectoral distribution of workers among Bolivian micro-enterprises is as follows: 20.3% industrial, 11.0% construction, 39.9% commerce, 26.4% services, and 2.6% others. The average number of years of schooling are considerably lower in the micro-enterprise sector than in the general population, they are 6.8 years in Bolivia's microfinance sector.

[5]The IADB defines a micro-enterprise as having no more than 10 workers and total assets below $US 20.000. Orlando and Pollack (2000), however, consider the number of workers only.

Figure 3.2: Unemployment rates by gender. Source: ILO.

3.2.2.2 Microfinance in Bolivia

Bolivia is deemed the most active microfinance market worldwide. There are many institutions, covering all big cities and most rural areas as well. While there is considerable competition in the big cities, the microfinance market in smaller towns and rural areas is less developed. There are currently three associations for microfinance institutions, ASOFIN (urban microfinance), CIPAME (support for microenterprises), and finrural (rural microfinance). The institutions covered in these associations had a combined portfolio of $US 2.8 Mio. in December 1990, growing to $US 382.7 Mio. in December 1999, which corresponds to roughly 0.5% of Bolivia's GNP. In June 2000 these institutions served 195,087 borrowers—only slightly fewer than the commercial banks with 218,956.

In December 1999, the largest institution was BancoSol with a portfolio of $US 82 Mio. The urban loans of all covered institutions totaled to $US 287 Mio. in December 1999. Their distribution is as follows: 51% to commerce, 14% to production, 17% to services, 9% to house improvements, 8% to consumption, and the rest to other destinations. Annual interest rates for loans in $US typically range from 25% to 35%, for loans in Bolivianos from 35% to 45% (nominal rates). Most institutions charge a flat rate commission between one and four per cent and loan sizes vary

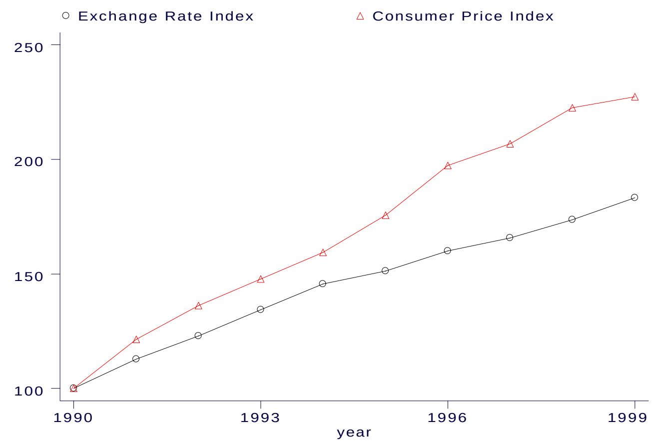

Figure 3.3: Inflation and exchange rate indices for Bolivia. Source: IADB.

between a minimum of $US 50 and a maximum of $US 120,000. Most institutions obtain funds from international aid institutions, e.g. in the form of long term loans with very low interest rates[6].

Bolivia's banking supervisory authority (Superintendencia de Bancos y Entidades Financieras, S.B.E.F.) provides credit records to all registered banks and private savings and loan companies. The institutions obtain credit information about all clients in their area including the amount of debts outstanding, the amount of guarantees outstanding, the amount and type of bad debts, and the name of the bank where the record originates from. While all registered institutions are participating, there is no information from Non-Governmental Organizations and other informal lenders (an inclusion of these institutions is planned).

The last two years have been characterized by challenging developments. In 1997, two banks began to offer consumer loans to the microfinance institutions' client base. These loans were disbursed based on credit scoring alone, no in-depth analysis of the clients' repayment capacity was made. Anecdotes about these lending practices abound: clients were asked whether they obtained a loan from one of the established microlenders. If so, they were offered a considerably higher consumer

[6]this data is from ASOFIN (2000), for a brief overview of the special characteristics of microfinance in Bolivia see also Rhyne (2001) and Inter-American Development Bank (1998)

loan without a further analysis of their repayment capacity.[7] Other clients were randomly assigned to a group. In these "group loans" the clients were co-debtors of all group members, frequently without knowing each other. Many clients borrowed large sums and eventually found themselves unable to fulfill their repayment obligations. This over-indebtedness of many clients leads to repayment problems of micro-loans taken earlier from other institutions, see also Rhyne (2001) and ASOFIN (July and December 1999) for a discussion of these issues.

The clients' repayment problems were reinforced by the economic crisis beginning in late 1998. Many clients faced severe drops in their incomes. As a consequence, repayment became even more difficult contributing to a severe rise in arrears, see table 3.3. The crisis not only affected the microlenders but the whole banking sector of the Bolivian economy. Between December 1999 and June 2000 the portfolio of the whole banking sector fell by 3.8% and portfolio in arrears for more than 30 days rose from 6.60% to 7.72%. Average annualized return on equity fell from 8.7% to 3.0% during the same time (source: banking supervisory authority S.B.E.F.). The crisis led to considerable job cuts. Many of the newly unemployed began to work in the informal sector, mostly as street vendors, leading to a rising competition which exacerbated the situation there. [8] [9]

3.3 Description of the Dataset

The data set covers the time from March 1992 to June 2000. It includes data from Caja Los Andes' predecessor Pro-Credito.[10] There is information on all six branches;

[7]These loans tend to have higher interest rates, covering part of the high costs of default.

[8]See various issues of "Nueva Economica" in July/August 2000.

[9]For more information on microfinance in Bolivia see Navajas, Schreiner, Meyer, Gonzalez-Vega, and Rogriguez-Meza (2000), who describe client characteristics, and Rhyne (2001) and Von Stauffenberg (2001) for an overview about recent developments. In an earlier study, Gonzalez-Vega, Meyer, Navajas, Schreiner, Rodriguez-Meza, and Monje (1996) analyze the client profiles of five relatively large microlenders in Bolivia with a special emphasis on their financial needs. Navajas, Conning, and Gonzalez-Vega (1999) analyze recent developments with a theoretical model, the focus being the competition between BancoSol, who uses mostly group loans, and Caja Los Andes, who focuses on individual loans.

[10]In personal conversation it was ensured that the operations and the client selection have not changed after the registration as a private bank. While the fraction of clients with relatively large loans has increased, Caja Los Andes still serves a large number of clients with very small loans.

Business sector	gender male	female	Total
Agriculture	74.74	25.26	100.00
	20.80	6.05	12.87
Commerce	29.35	70.65	100.00
	29.29	60.67	46.15
Stockbreeding	48.83	51.17	100.00
	0.65	0.59	0.62
Production	53.35	46.65	100.00
	26.52	19.96	22.99
Service	60.57	39.43	100.00
	22.74	12.74	17.36

Table 3.5: Loans disbursed between 01/1999 and 06/2000 by gender and business sector (in %).

their different sizes are documented in tables 3.1, 3.21, and 3.22. The branch in La Paz includes El Alto, a city with over a million inhabitants and a very large client base. Operations in Santa Cruz, on the other hand, have only begun recently and Caja Los Andes tries to establish a foothold in its highly competitive market.

The data set includes information on the clients, such as gender, age, and civil status. The loan data includes information on the amount applied for, the amount granted, terms and conditions, and repayment behavior, among many other variables. Furthermore there are estimated balance sheets for most clients. Whenever a client applies for a loan, a loan officer goes to the business and estimates the assets, liabilities, sales revenues, expenses, and number of employees, and more. This information is gathered irregularly. Clients with a good repayment performance eventually obtain an automatic credit line and can take out new loans without a detailed business check. The following sections provide a description of some characteristics of these data.

3.3.1 Clients

Of all loans disbursed between 01/1999 and 06/2000, 46% went to men and 54% to women. The distribution varies considerably between sectors: While of all loans in agriculture only 26% went to women, 71% of all loans to the commerce sector went to women, see also table 3.5.[11] The clients' civil status is frequently used as an indicator of stability by the loan officers. Between 01/1999 and 06/2000 most loans went to married people (67%). Again, there are differences between the business sectors. 74% of all agricultural loans are disbursed to married clients while the fraction of married clients in commerce and service loans is lower with 65% and 62%, respectively. Regional differences abound as well, the fraction of married clients in Santa Cruz being 46% only. While between 01/1999 and 06/2000 15% of the active clients lived alone, 22% lived in households with 10 or more persons. Households in Santa Cruz tend to be relatively small, with 19% single households and 18% living in households with 10 or more people.

What is the age of the clients taking their first loans from Caja Los Andes? From January to June 2000 24% of the new customers were in their twenties, while most clients were between thirty and forty (33%). 25% were between forty and fifty, 13% between fifty and sixty, and 4% were older than sixty when taking their first loan.

Due to the high competition in the Bolivian microfinance market, many clients have multiple loans at different institutions. Of all clients who had an outstanding loan at Caja Los Andes between January and June 2000, 34% had loans with other regulated institutions as well.[12] For clients with loans between $US 5,000 and $US 10,000 the number is 49%, for even larger loans the number is 53%. In other words, clients with larger loans tend to have loans from other (regulated) sources as well while clients with smaller loans more often either do not want other loans or do not have access to these sources; see also section 3.3.4 .

Many clients take repeat loans from Caja Los Andes. Of all clients taking their first loan in 1992, 88% took a second loan at some later point in time, 79% took a

[11] If no other sources are mentioned, the information provided in the tables is based on calculations from Caja Los Andes' data set.

[12] Source: calculations based on the credit information of the S.B.E.F. These numbers include loans from other regulated institutions only.

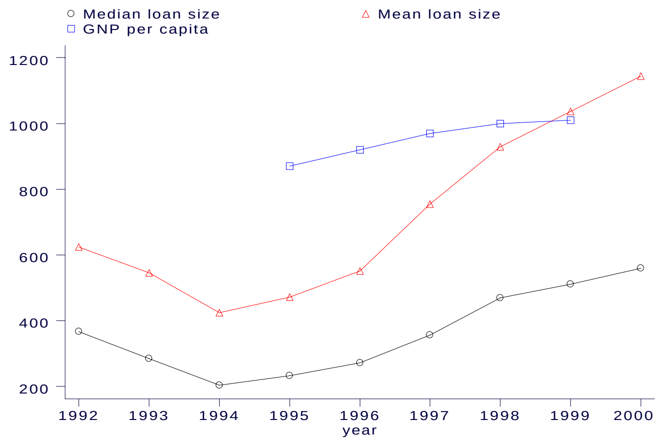

Figure 3.4: Median and mean loan sizes compared to GNP per capita. All values are in 1992 $US, source: own calculations and Worldbank (2000a).

third loan and 70% took four or more loans from Caja Los Andes until June 2000. Over time, these numbers have decreased slightly. Of all clients who took their first loan in 1995 only 81% took a second loan and, so far, only 55% took a fourth loan, see table 3.30 in the appendix.

3.3.2 Loans

The number of loans disbursed has increased considerably during the first years of operation from 3,989 in 1992 to 39,377 in 1996. In recent years, the number of loans disbursed has decreased slightly to 35,089 in 1999, see table 3.21 for an overview. Although the number of disbursed loans has decreased, the total amount outstanding has increased each year, owing to an increase in the average loan life and loan size. The number of loans outstanding has increased from 1,183 at December 31st 1992 to 35,171 at June 30th 2000, see table 3.6.

The development of loan sizes has followed two different phases. Initially (1992), 62% of all loans were below $US 500. The fraction of these small loans increased until 1995. Since 1995, however, loan sizes have increased, which can best be seen in table 3.4. The average loan size has increased from $US 472 in 1995 to $US 1,144 in June 2000. The 10th percentile has increased from $US 77 to $US 170 over the same

Date	Branch of observation						
Outstand.	cbb	lpb	scz	sre	tdd	tja	Total
31.12.92		1,183					1,183
31.12.93		3,299					3,299
31.12.94		7,435		294			7,729
31.12.95	1,278	10,779		1,193		592	13,842
31.12.96	2,493	15,661		1,840		1,281	21,275
31.12.97	2,759	17,154		2,198	514	1,775	24,400
31.12.98	3,118	19,635		2,561	1,197	2,780	29,291
31.12.99	4,476	20,226	431	3,151	1,593	3,294	33,171
31.06.00	4,845	21,210	857	3,261	1,487	3,511	35,171

Table 3.6: Number of outstanding loans by year and branch.

time period and the 90th from \$US 966 to \$US 2,578 (all values are in 1992 \$US). When compared to GNP per capita, the mean loan size has increased considerably stronger, see also figure 3.4.

For very small loans, not only the fraction but also the number of these loans has decreased over time. While in 1995 23,937 loans with less than \$US 500 (1992 values) were distributed, constituting 76.68% of loan contracts, this number has fallen to 17,242 (49.54%) in 1999. The fraction of loans between \$US 500 and \$US 5,000, on the other hand, has increased considerably from 22.64% in 1995 to 47.56% in 1999. The fraction of loans larger than \$US 5,000 has increased from 0.69% to 2.89%. For further details see table 3.25 in the appendix. The change in the distribution of loans disbursed is amplified when considering the amounts of all disbursed loans in a certain range, see table 3.27. The amounts disbursed in loans of less than \$US 500 have decreased after 1996 while the amounts disbursed in larger loans have increased.

To see whether this change in loan sizes reflects a change in Caja Los Andes' policy or changing demands for loans, consider the amounts applied for. Over all observations, the median relation between amount approved and amount applied for is 97%, while 10% of all loan amounts are less or equal to 50% of the applied amounts. The median has increased from 0.75 in 1993 to 1 in 1999/2000, the number is lower for first loans and higher for consecutive loans. In other words, clients are

	Business sector					
gender	Agriculture	Commerce	Stockbreed.	Production	Service	Total
male	157	123	170	109	123	122
	424	512	526	389	468	440
	1,709	6,328	2,633	3,485	4,039	4,348
female	146	90	166	90	85	90
	405	327	514	301	302	320
	1,399	2,608	1,733	2,148	2,578	2,434
Total	152	96	169	99	97	99
	420	361	516	343	391	362
	1,709	3,517	2,186	2,709	3,417	3,077

Table 3.7: 10th percentile, median, and 90th percentile of loan sizes by gender and business sector. January 1999 to June 2000, amounts are in 1992 $US.

more likely to obtain the desired amount when they apply in later years and/or for repeat loans.

Loan sizes vary considerably by business sector and by gender. The largest loans are distributed in the stockbreeding and agriculture sectors with medians of $US 516 and $US 420 for 1999/June 2000, respectively. Male clients take out considerably larger loans than female clients, the medians being $US 440 and $US 320, respectively, see table 3.7. The differences are most pronounced in the non-agricultural sectors. Loans tend to increase as the client takes repeat loans. The median loan size growth rate is 41%. For 1999/June 2000, 1.46% of first loans (new clients) were above $US 5,000, 2.48% of second loans, 3.92% of third loans, and 5.18% of fourth or higher loans. Again, the values are in 1992 $US.

Interest rates vary. For loans denominated in $US in 01/1999 to 06/2000, 248 (1%) had a monthly interest rate of less than 2% (nominal rates). The bulk of the loans (76%) had a monthly interest rate between 2.5% and 3.5%. For loans denominated in Bolivianos, the rate is between 3% and 3.5%, for loans denominated in inflation adjusted Bolivianos, the rate is either 2% or 2.5%. The link between interest and loan sizes is unclear. While loans below $US 500 and loans above $US 5000 tend to have relatively low rates, rates are higher for the intermediate range, see table 3.26 in the appendix.

Year loan	Currency code		
is taken	US Dollar	Boliviano	infl. adj.
1992	0.00	98.34	1.66
1993	0.00	83.51	16.49
1994	0.00	97.85	2.15
1995	1.22	96.21	2.57
1996	12.66	84.75	2.59
1997	27.39	67.67	4.94
1998	33.04	59.12	7.84
1999	37.28	55.40	7.32
June 2000	43.16	54.35	2.50

Table 3.8: Distribution of currencies of disbursed loans over time in percentage values.

Initially, most loans (98%) were disbursed in Bolivianos with a small fraction of loans in inflation adjusted Bolivianos, see table 3.8. Since 1995 Caja Los Andes offers loans denominated in $US, in the first half of 2000 these loans made up 43% of all loans disbursed. Larger loans tend to be denominated in $US more frequently, see table 3.28 in the appendix.

Over the years of operation, the average duration of loans has increased considerably from 80 days in 1992 to 531 days in 2000, see table 3.29. The service sector on average has the loans with the longest duration (502 days in 1999) while the agriculture sector has the shortest loans (360 in 1999). The duration increases slightly for repeat loans. Between 01/1999 and 06/2000, the average duration for first loans was 420 days, 464 days for second loans, and 515 days for third or higher loans. The distribution between new loans, second, third and further loans has been relatively stable over time. Between 30% and 40% of all loans disbursed in any year are new loans, roughly 20% are second loans, and between 40% and 50% are third or higher loans.

	Number of days late this payment				
year	0	1-2	3-9	10-29	≥30
1996	75.12	12.09	9.75	2.51	0.53
1997	68.01	13.05	12.66	4.86	1.43
1998	63.90	12.87	14.74	6.18	2.31
1999	75.46	6.35	9.19	5.38	3.62
2000	75.45	4.73	8.17	6.47	5.18

Table 3.9: Fraction of payments with arrears over time (in %).

3.3.3 Repayment Behavior

The data contains detailed information about the clients' payments. In the first years of operation (1992 to 1994) only the maximum number of days in arrears has been recorded in most cases (days in arrears correspond to days overdue). In later years the data contains information about the exact date of each payment and the number of days each payment was late or early. Table 3.9 displays the fraction of clients with arrears of a given size over time. There is a strong increase in the fraction of payments with arrears of 30 days or more from 0.53% in 1996 to 5.18% in 2000.

Repayment behavior varies by loan size. It is worst for loans of a size between $US 1,000 and $US 10,000 (values are in 1992 $US) where the fraction of all payments that were at least 30 days late rose from 0.38% in 1996 to 6.00% in the first half of 2000 (see table 3.31 in the appendix). For loans smaller than $US 100 this fraction has been lower in all years and it rose from 0.33% to 3.16%. When comparing different sectors we find that the fraction of payments at least 30 days late is especially high in the agricultural sectors (8.30% in the first half of 2000) and lower in the other sectors.

After the severe repayment problems beginning in 1998 there is a strong increase in the fraction of payments without arrears and a decrease in the fraction of payments a few (one to nine) days late. Between 1998 and 1999 the former rose from 64% to 76% and the latter declined from 27% to 15%. These numbers reflect Caja Los Andes' increased concern about late repayments and the ensuing rise in repayment enforcement. Table **??** shows the number of payments 1 or 2 days late relative to

	Branch of observation					
year	cbb	lpb	scz	sre	tdd	tja
1995	0.15	0.12		0.17		0.12
1996	0.16	0.17		0.19		0.13
1997	0.21	0.20		0.20	0.11	0.15
1998	0.26	0.20		0.19	0.15	0.18
1999	0.12	0.07	0.19	0.22	0.20	0.07
2000	0.07	0.03	0.24	0.21	0.18	0.08

Table 3.10: The ratio of the number of payments one or two days late to punctual payments as a proxy for the enforcement of punctual repayment.

punctual payments for each branch. A stronger repayment enforcement is visible in La Paz, Cochabamba, and Tarija.

3.3.4 Businesses

Whenever a client applies for a new loan, balance information is estimated by the loan officer. Since 1995, there have been roughly 30.000 balance observations per year. Divided by the number of active clients given in table 3.1 this yields close to one observation per active client per year, the number decreasing slightly.

For all clients active between January 1999 and June 2000, 98% owned their businesses. 0.74% had a formal registration; all of these businesses are non-agricultural. Half of all businesses were founded before or in 1993, 25% before or in 1987. The highest fraction of old businesses is found in agriculture and stockbreeding, with 6% and 11% founded before 1960, respectively. The youngest sector are the services, with 42% founded between 1996 and 2000, see table 3.12.

Most businesses have no employees (95% for business information between January 1999 and June 2000). Businesses with at least one employee are most frequently found in the production sector. Of all businesses in 01/1999 to 06/2000, 11% of all production businesses, 7% of all service businesses, and 3% of all commerce businesses had at least one employee.

During the same time, median assets held were $US 2,843, see table 3.13. The stockbreeding sector has the highest assets with a median of $US 10,282 and the

	Branch of observation						
Year	cbb	lpb	scz	sre	tdd	tja	Total
1992		1,332					1,332
1993		8,544					8,544
1994		21,756		397			22,153
1995	2,691	30,022		3,754		1,400	37,867
1996	6,596	33,665		4,224		3,468	47,953
1997	5,909	22,054		3,925	788	3,600	36,276
1998	5,778	19,244		4,038	2,010	5,371	36,441
1999	4,679	24,665	620	4,456	2,383	2,503	39,306
June 2000	1,455	11,822	684	2,103	991	654	17,709
Total	27,108	173,104	1,304	22,897	6,172	16,996	247,581

Table 3.11: Number of balance observations by year and branch.

10th percentile as high as $US 33,504. The production sector has the lowest median assets with $US 2,105. Assets in the commerce sector are relatively low as well with a median of $US 3,409, compared to $US 5,836 in agriculture and $US 3,409 in services. Asset holdings differ across regions with the highest median asset holdings in Cochabamba of 3,989. Between 01/1999 and 06/2000 businesses classified as small enterprises had median assets of $US 41,641 whereas micro-enterprises had median assets of $US 2,475. Women had median assets of $US 2,178 compared to $US 3,592 for men. Over time, median assets over all branches have increased from $US 1,175 in 1992 to $US 3,048 in 2000. Again, all values are in 1992 $US.

The data also contains information about liabilities which makes it possible to

year of business foundation	Freq.	Percent
before 1960	713	1.35
sixties	1,390	2.63
seventies	4,078	7.71
eighties	10,529	19.91
90 to 95	18,015	34.06
96 to 2000	18,163	34.34

Table 3.12: Year of business foundation of all clients active between January 1999 and June 2000.

Business Sector	Branch of observation cbb	lpb	scz	sre	tdd	tja	Total
Agriculture	3,414	1,035		1,304	1,920	1,236	1,575
	8,528	2,832		4,671	11,822	3,692	5,836
	27,159	8,884		15,401	20,636	13,794	20,264
Commerce	713	402	792	351	738	605	432
	3,637	2,193	3,703	1,987	3,485	2,655	2,473
	15,025	18,475	16,536	15,470	15,902	14,379	17,268
Stockbreeding	3,259	944		271			2,961
	10,659	37,107		1,385			10,282
	33,504	73,270		5,870			33,504
Production	1,084	491	324	318	620	494	477
	4,252	2,023	1,908	1,390	2,296	1,890	2,105
	17,894	11,774	11,162	9,894	11,147	11,533	12,094
Service	606	444	663	799	656	777	553
	3,419	3,138	3,224	4,611	2,666	3,530	3,409
	16,694	14,309	18,780	21,510	12,391	20,392	16,060
Total	1,186	442	663	424	699	675	488
	5,989	2,316	3,188	3,089	3,119	3,031	2,843
	22,028	15,147	15,893	16,643	14,928	15,221	16,245

Table 3.13: 10th percentile, median, and 90th percentile of assets by branch and business sector for observations between January 1999 and June 2000. Values are in 1992 $US.

Year	1992	1993	1994	1995	1996	1997	1998	1999	2000
Trade Credits	20.3	17.9	13.7	9.7	5.9	5.7	5.4	6.1	6.4
Other Loans	10.4	6.6	7.3	7.5	10.5	8.7	10.2	23.1	27.0

Table 3.14: Fraction of clients with trade credits and loans from other sources (in %).

calculate equity. However, this information is not very reliable for numerous reasons: firstly, all loans from informal lenders or relatives have to be self-reported and it might not lie in the households' interest to report the correct amount. Secondly, data from loans from other regulated institutions is available from the superintendency, but the data is frequently either not received in time or not recorded correctly. For these reasons we will consider some brief statistics only. Median equity in 01/1999 to 06/2000 was $US 2,647, with the 10th percentile $US 455, the 90th percentile $US 15,426. Median equity is highest in the stockbreeding and agricultural sectors with $US 11,461 and $US 5,906, respectively. Median annualized equity growth equals 6.1% and tends to be slightly higher in the commerce sector than in the production sector.

More than 81% of all balance information contain neither trade credits nor loans from other sources. The importance of trade credits has decreased considerably over time. While in 1992 20.3% of all businesses obtained trade credits, only 6.4% did so in 2000, see table 3.14. This decline is similar in all branches and for all business sectors. The number of businesses with loans from other sources, however, has increased from a low of 6.6% in 1993 to 27% in 2000, indicating the increasing availability of loans. These numbers are particularly high in Cochabamba and Santa Cruz, with 28% and 50%, respectively. For small businesses with assets below $US 1,000 the fraction of businesses with trade credits has been smaller (16% in 1992 and 3% in 2000). In contrast to larger businesses, the fraction of these businesses having loans from other sources has decreased from 1992 (7.2%) to 1998 (2.4%) and increased again to 8.5% in 2000. The better availability of loans for larger businesses is similarly evident when regarding the average liability over assets. While it is 0.04 for businesses with assets between $US 1,000 and 5,000, it is 0.15 for businesses with assets above \$US 50,000. Again, all values are in 1992 \$US.

The balance data also contains information on business and non-business income. Since the distinction between both is difficult for agricultural businesses, there is only few and not very reliable income information for the respective sectors. In the following paragraphs, we focus on the non-agricultural sectors—commerce, production, and services.

Median monthly business income between 01/1999 and 06/2000 was \$US 186, see table 3.15. It was highest in Santa Cruz and Cochabamba, with a median of \$US 210 and \$US 211, respectively. When comparing different business sectors one finds that the median income is highest in services with \$US 206. Between 01/1999 and 06/2000, small enterprises had a median business income of \$US 1,043 while micro-enterprises had a median business income of 181. The median for males was \$US 219 while the median for females was \$US 160.

The data also includes information about non-business income which allows to calculate the total income of a client. For 1999/June 2000 median total monthly income was \$US 246, mean total monthly income \$US 335. Over time, the median has decreased in real terms from \$US 333 in 1992 to \$US 241 in 2000. Again there are differences between the median income of small enterprises (\$US 1,093) and micro-enterprises (\$US 241), between men (\$US 262) and women (\$US 236). Values are in 1992 \$US. For all observations the median share of business income in total income is 0.85. It is highest in the production sector, where more than half of the clients obtain all their income from their businesses.

We now can calculate the return on assets for each business. The median monthly return on assets between 01/1999 and 06/2000 is 7% for the commerce sector, 9% for production activities, and 6% for services. While the 10th percentile is 2% for these three branches, the 90th percentile is between 23% and 25%. When distinguishing between small and micro-enterprises we find that between 01/1999 and 06/2000 small enterprises had a median return on assets of 2.4% while micro-enterprises had a median return on assets of 7.4%. The median for males was 6.4% and the median for females was 7.5%, corresponding to the differences in assets.

Since we have multiple observations of the same clients, we can calculate annu-

Business	Branch of observation						
sector	cbb	lpb	scz	sre	tdd	tja	Total
Commerce	60	47	58	35	67	64	48
	191	173	206	146	199	187	175
	477	583	455	486	512	519	549
Production	90	69	30	17	80	74	64
	254	191	197	122	181	204	190
	555	487	410	346	432	447	479
Service	81	69	99	59	81	76	71
	212	211	233	188	182	219	206
	507	501	476	436	446	544	483
Total	67	56	61	36	73	69	55
	211	187	210	155	191	201	186
	511	537	458	441	483	519	514

Table 3.15: 10th percentile, median, mean, and 90th percentile of monthly business income by branch and business sector for observations between January 1999 and June 2000.

alized growth rates e.g. for assets using

$$growth_{assets} = \left(\frac{assets_t}{assets_{t-1}}\right)^{\frac{day_t - day_{t-1}}{365}} - 1 \quad .$$

We find that between 01/1999 and 06/2000 median annualized asset growth is 0.13. Asset growth is highest in commerce with a median of 0.16 and in the Trinidad branch with a median of 0.29. From 1993 to 1999 median asset growth has increased from 0.02 to 0.15, falling to 0.07 in the first half of 2000. Asset growth is highest for smaller businesses. If assets are below $US 1,000, asset growth in the non-agricultural sectors has risen from a median of 0.04 in 1994 to 0.27 in 1999. Correspondingly, asset growth is higher for women and micro-enterprises than for men and small enterprises.

For January to June 2000 the median annualized growth in business income was -0.02 in all non-agricultural sectors. The 10th percentile was -0.64, the 90th percentile 1.08. The relatively low numbers reflect the economic crisis. The branches which were hit the most were Tarija, with a median growth in business income of -0.07, and Sucre with a median of -0.04. Small enterprises seem to have suffered the most

Business sector	Branch of observation cbb	lpb	scz	sre	tdd	tja	Total
Commerce	0.014	0.017	0.011	0.012	0.014	0.020	0.016
	0.053	0.075	0.061	0.077	0.057	0.071	0.071
	0.185	0.243	0.186	0.249	0.193	0.205	0.230
Production	0.016	0.023	0.026	0.005	0.018	0.023	0.021
	0.058	0.093	0.094	0.085	0.075	0.106	0.089
	0.170	0.258	0.256	0.288	0.262	0.311	0.257
Service	0.016	0.019	0.016	0.009	0.018	0.016	0.016
	0.062	0.069	0.063	0.040	0.069	0.057	0.062
	0.241	0.283	0.257	0.164	0.220	0.192	0.248

Table 3.16: 10th percentile, median, and 90th percentile of monthly return on assets by branch and business sector for observations between January 1999 and June 2000.

with a median growth of -0.06 in 01/1999 to 06/2000, women are slightly better off than men with a median of 0.005 (compared to -0.005 for men).

Median growth in total income was -0.01 in 2000 (0.01 in 1999) while it was 0.04 in 1997 and 1998. Again, small enterprises have lost the most with a median of -0.05 in 01/1999 to 06/2000. Median sales growth has been zero in the first half of 2000, down from 0.03 in 1999 and 0.06 in 1998. Again, micro-enterprises are better off with median sales growth of 0.02 in 01/1999 to 06/2000; so are women with 0.02 compared to men with 0.01.

3.4 Hypotheses

3.4.1 Return on Assets Decreases for Clients

This section examines the following hypothesis:

A successful provision of microfinance implies an alleviation of credit constraints for micro-enterprises and it allows the business to increase the capital level used in production. Assuming decreasing returns to scale and a sub-optimal initial level of capital, the increase in assets moves the return on assets downwards closer to the market interest rate. This effect should be more pronounced for very small busi-

Year of observation	Year of first observation 1992	1993	1994	1995	1996	1997	1998	1999	2000
1992	877								
1993	2,172	1,429							
1994	2,058	1,822	884						
1995	2,734	2,508	1,284	1,006					
1996	1,303	2,913	1,649	1,327	827				
1997	1,886	3,826	2,030	1,908	1,156	945			
1998	3,429	3,935	2,838	2,600	1,707	1,343	1,067		
1999	4,595	5,574	4,493	4,212	2,881	2,627	1,685	1,881	
2000	1,998	5,320	4,236	4,282	3,479	3,128	2,395	2,379	2,346

Table 3.17: Median assets by year of first observation (cohort), commerce sector. Values are in 1992 $US.

nesses, since they face the greatest credit constraints.

Consider, for example, a Cobb-Douglas production function $Y = AL^{\alpha}K^{1-\alpha}$ with a marginal return on capital of $MRK = A(1-\alpha)\left(\frac{L}{K}\right)^{\alpha}$. If the business does not face credit constraints, the optimal level of capital is such that the marginal return on capital equals the market interest rate. Credit constraints reduce the level of capital used. All else equal, the marginal return on capital is higher for credit constrained firms since MRK decreases in K. Put differently, the alleviation of binding credit constraints leads to a decrease in the return on capital. In the following paragraphs, we proceed in two steps. Firstly, we show that capital usage increases for clients who take out repeat loans, where we approximate capital by the sum of all assets. Secondly, we show that the return on capital decreases over time for these clients.[13]

Since the amount of assets held varies considerably between sectors, the following analysis focuses on the commerce sector which is where most of Caja Los Andes' clients have their businesses. The developments in the other sectors are similar. Table 3.17 shows that initial assets in the commerce sector have risen from a median of $US 877 in 1992 to a median of $US 2,346 in 2000, again all values are in 1992 $US. When regarding the median assets by cohort we also find a rise in most cases.

[13]The amount of labor stays constant over time for the largest part of the sample. A decreasing productivity parameter A would also lead to a decline of the return on capital, e.g. during the economic crisis. To avoid attributing a productivity based decline in the MRK to better capital supply we not only consider the return on assets but also the development of assets over time.

Year of observ.	Year of first observation (cohort) 1992	1993	1994	1995	1996	1997	1998	1999	2000
1992	0.356								
1993	0.278	0.278							
1994	0.186	0.208	0.225						
1995	0.134	0.154	0.178	0.210					
1996	0.131	0.135	0.149	0.169	0.196				
1997	0.137	0.124	0.137	0.140	0.167	0.195			
1998	0.104	0.121	0.119	0.122	0.139	0.166	0.196		
1999	0.058	0.077	0.082	0.086	0.103	0.117	0.145	0.177	
2000	0.101	0.060	0.067	0.076	0.088	0.092	0.114	0.137	0.181

Table 3.18: Monthly median return on assets by year and cohort for clients with initial assets below $US 1,000. Commerce sector.

For clients with their first balance observation in 1994, for example, median assets have risen from $US 884 in 1994 to $US 4,236 in 2000 (+380%). This strong increase in assets supports our hypothesis.[14]

Regarding the return on assets, we find a considerable decrease from a median monthly return on assets of 20.9% in 1992 to 6.6% in 2000.[15] This decrease has two possible sources. Firstly, capital supply for small and micro-enterprises has improved in the Bolivian economy, see also section 3.2.2.2. Thus, even new clients should have a higher leverage than they used to in earlier years. Secondly, clients taking repeat loans from Caja Los Andes, mostly of growing size, can increase their assets.

To separate these effects, we can compare the median return on assets of different cohorts. Consider the commerce sector and micro-enterprises with initial assets below $US 1,000 as presented in table 3.18. The median return on assets for new clients has decreased from 36% in 1992 to 18% in 2000. Again we also find a decrease

[14]While we do not control for selection bias at this point, we can rule out that the increase in assets solely stems from a dropout of clients with low assets. When we make the same cohort analysis restricted to clients which we observe at least three times, for example, the rise in assets is similar.

[15]One can think of other possible explanations of the declining return on assets, e.g. the recession beginning in 1998 or increasing competition from the rising number of urban poor. While we can rule out the recession as the main cause since it begins only in late 1998 while return on assets declines in earlier years as well, a more comprehensive econometric analysis is needed to single out other factors.

of the return on assets for each cohort. Clients taking their first loan in 1994, for example, had a median monthly return on assets of 22.5%, declining to 6.7% in 2000. In addition, we find that clients who have been with Caja Los Andes for a longer time (earlier cohorts) have a lower return on assets than new clients. In 1998, for example, new clients had a median return on assets of 19.6%, while clients with their first observation in 1994 had a median return on assets of 11.9%. These effects indicate a positive influence of the loans from Caja Los Andes on the clients' use of assets.

While the numbers presented above show a rise in assets and a decline in return on assets consistent with our hypothesis, selection bias is a considerable problem here. We have multiple observations only of those clients who take repeat loans. These clients could well be more successful than others. These issues can be tackled in a more comprehensive regression analysis which includes the estimation of survival probabilities. This is left for future work (chapter 4).

3.4.2 Business Income Increases for Clients

This section examines the following hypothesis:

The provision of microfinance should lead to a rise in business income through a) a reduction in interest costs and b) the alleviation of capital constraints. Thus, business income should be higher for clients after their second, third etc. loans. Again, this effect should be larger for smaller businesses since they face larger credit constraints.

To examine this hypothesis consider the median monthly business income for different cohorts as shown in table 3.19. The table again shows two effects. Firstly, median business income increases over time e.g. from $US 160 for clients taking their first loan in 1994 to $US 284 in 1998 (+78%). Secondly, for any given year, clients who are longer with Caja Los Andes than others (earlier cohorts) tend to have higher business income as well. For observations made in 1997, for example, new clients had a median business income of $US 139, clients with a first loan in 1996 had a median of 162, and clients with a first loan in 1993 had a median of 283.

Year of observ.	Year of first observation (cohort) 1992	1993	1994	1995	1996	1997	1998	1999	2000
1992	226								
1993	349	241							
1994	196	251	160						
1995	249	314	187	156					
1996	195	296	208	177	133				
1997	238	283	241	213	162	139			
1998	310	385	284	247	194	176	156		
1999	378	294	248	243	214	200	170	162	
2000	538	324	245	249	181	190	171	167	157

Table 3.19: Median monthly business income by year and cohort for clients in the commerce sector. Values are in 1992 $US.

In 1999 and 2000 median business income decreases for most cohorts, coinciding with the recession beginning in late 1998. With the exception of this decrease the development of median business income of different cohorts over time is consistent with our hypothesis since it shows that median business income grows over time.

How does this compare to the development of smaller businesses? Table 3.20 lists median business income for clients with initial assets below $US 1,000 by the year of the first client observation (cohort). The table shows the same two effects. Firstly, business income increases over time e.g. from $US 107 for clients taking their first loan in 1994 to $US 196 in 1999. Secondly, for any given year clients who are longer with Caja Los Andes than others (earlier cohorts) tend to have higher business income as well. For observations made in 1997, for example, new clients had a median business income of $US 93, clients with a first loan in 1995 had a median of 140, and clients with a first loan in 1992 had a median of 158. When we compare the change in business income of the low asset group (table 3.20) with the total sample (table 3.19) we find that the relative increase in median business income is smaller for the low asset group. That is, the median business income of the 1993 cohort, for example, has grown 60% until 1998 for all commerce businesses (385 compared to 241) while it has grown 33% only for the low asset group (194 compared to 146). This could indicate the existence of scale effects and stands in contrast to our hypothesis.

Year of	Year of first observation (cohort)								
observ.	1992	1993	1994	1995	1996	1997	1998	1999	2000
1992	150								
1993	164	146							
1994	190	155	107						
1995	184	198	134	100					
1996	180	160	146	118	92				
1997	158	156	152	140	113	93			
1998	233	194	166	150	129	112	97		
1999	270	199	196	156	140	117	110	89	
2000	134	235	145	153	130	119	103	94	84

Table 3.20: Median monthly business income by year and cohort for clients with initial assets below $US 1,000, commerce sector. Values are in 1992 $US.

3.5 Appendix

A short note on the tables Branches are abbreviated as follows: cbb = Cochabamba, lpb = La Paz, scz = Santa Cruz, sre = Sucre, tdd = Trinidad, and tja = Tarija. Unless mentioned otherwise, all tables are based on own calculations from Caja Los Andes' data set and all values are in 1992 $US.

	Branch of observation						
Year	cbb	lpb	scz	sre	tdd	tja	Total
1992		3,849					3,849
1993		7,621					7,621
1994		19,078		324			19,402
1995	2,393	22,879		3,108		1,122	29,502
1996	5,195	25,614		3,326		2,597	36,732
1997	4,537	24,657		3,066	570	2,513	35,343
1998	3,978	22,263		3,027	1,518	3,241	34,027
1999	5,019	18,269	452	3,330	1,647	3,374	32,091
June 2000	3,834	9,354	521	1,508	590	1,942	17,749

Table 3.21: Number of loans disbursed by year and branch.

Year of	Branch of observation						
first loan	cbb	lpb	scz	sre	tdd	tja	Total
1992		1,518					1,518
1993		2,843					2,843
1994		6,344		306			6,650
1995	1,511	5,918		1,373		724	9,526
1996	2,285	7,765		1,151		1,150	12,351
1997	1,462	6,800		912	534	969	10,677
1998	1,318	6,514		953	909	1,501	11,195
1999	2,404	5,661	436	1,317	762	1,251	11,831
June 2000	1,928	3,086	460	581	146	709	6,910

Table 3.22: Number of new clients by year and branch.

Sector	number of contracts	%	amount	%
Service	3,797	17.3	41,151,615.18	18.1
Commerce	10,854	49.3	124,021,906.33	54.4
Production	6,431	29.2	57,663,027.46	25.3
Agriculture	876	4.0	4,975,832.79	2.2
Stockbreeding	1	0.0	3,000.00	0.0
others	35	0.2	63,137.70	0.0
total	21,994	100.0	227,878,519.47	100.0

Type of guarantee	number of contracts	%	amount	%
chattel	8,545	38.9	67,824,600.68	29.8
mixed	65	0.3	7,187,885.56	3.2
chattel and personal	13,249	60.2	134,675,868.72	59.1
mortgage and personal	1	0.0	372,000.00	0.2
mortgage and chattel	134	0.6	17,818,164.50	7.8
total	21,994	100.0	227,878,519.47	100.0

payment frequency	number of contracts	%	amount	%
weekly	75	0.3	500,245.59	0.2
bi-weekly	752	3.4	3,893,133.42	1.7
monthly	16,301	74.1	151,959,623.49	66.7
three-monthly	6	0.0	1,333,389.67	0.6
two-monthly	2	0.0	44,600.00	0.0
irregular	4,858	22.1	70,147,527.30	30.8
total	21,994	100.0	227,878,519.47	100.0

loan destination	number of contracts	%	amount	%
working capital	17,448	79.3	162,074,853.63	71.1
fixed capital	3,513	16.0	54,450,521.03	23.9
mixed (work.+ fixed)	941	4.3	10,981,454.93	4.8
house improvement	17	0.1	160,305.38	0.1
consumption	58	0.3	131,331.50	0.1
freely disposable	17	0.1	80,053.00	0.0
total	21,994	100.0	227,878,519.47	100.0

Table 3.23: Characteristics of loans outstanding on July 31st, 2000, in La Paz. Source: Caja Los Andes.

duration	number of contracts	%	amount	%
up to 3 months	69	0.3	524,038.40	0.2
4 to 6 months	296	1.3	1,068,495.15	0.5
7 to 9 months	755	3.4	2,149,446.40	0.9
10 to 12 months	4,699	21.4	17,888,516.27	7.9
13 to 18 months	6,621	30.1	44,902,951.67	19.7
19 to 24 months	5,261	23.9	64,600,597.35	28.3
more than 24 months	4,293	19.5	96,744,474.24	42.5
total	21,994	100.0	227,878,519.47	100.0

sex	number of contracts	%	amount	%
females	12,456	56.6	113,007,504.50	49.6
males	9,538	43.4	114,871,014.96	50.4
total	21,994	100.0	227,878,519.47	100.0

type	number of contracts	%	amount	%
new	6,600	30.0	51,416,220.23	22.6
recurrent	15,394	70.0	176,462,299.23	77.4
Total	21,994	100.0	227,878,519.47	100.0

days in arrears	number of contracts	%	amount	%
no arrears	19,375	88.1	203,058,952.78	89.1
1 to 10	464	2.1	4,264,664.10	1.9
11 to 20	381	1.7	3,715,814.00	1.6
21 to 30	394	1.8	3,706,097.89	1.6
31 to 90	446	2.0	4,563,210.42	2.0
more than 90	934	4.2	8,569,780.28	3.8
total	21,994	100.0	227,878,519.47	100.0

Table 3.24: Characteristics of loans outstanding on July 31st, 2000, in La Paz, continued. Source: Caja Los Andes

Year	Sizes of disbursed loans < 500	500-5000	> 5000	Total
1992	2,473	1,474	18	3,965
	(62.37%)	(37.18%)	(0.45%)	(100.00%)
1993	5,474	2,305	39	7,818
	(70.02%)	(29.48%)	(0.50%)	(100.00%)
1994	15,429	4,603	80	20,112
	(76.72%)	(22.89%)	(0.40%)	(100.00%)
1995	23,937	7,067	214	31,218
	(76.68%)	(22.64%)	(0.69%)	(100.00%)
1996	27,526	11,020	367	38,913
	(70.74%)	(28.32%)	(0.94%)	(100.00%)
1997	22,487	14,249	692	37,428
	(60.08%)	(38.07%)	(1.85%)	(100.00%)
1998	20,165	15,258	912	36,335
	(55.50%)	(41.99%)	(2.51%)	(100.00%)
1999	17,242	16,554	1,007	34,803
	(49.54%)	(47.56%)	(2.89%)	(100.00%)
June 2000	8,519	9,221	637	18,377
	(46.36%)	(50.18%)	(3.47%)	(100.00%)

Table 3.25: Development of loan sizes over time by number of disbursed loans.

Loan sizes	Interest rates < 1.5	1.5 to 2	2 to 2.5	2.5 to 3	Total
< 500		2.21	33.73	64.06	100.00
500-5000	0.01	0.12	19.26	80.60	100.00
> 5000	4.22	8.13	37.65	50.00	100.00
Total	0.34	0.97	22.20	76.48	100.00

Table 3.26: Distribution of interest rates for various loan sizes for loans denominated in \$US between 01/1999 and 06/2000 (in %).

Year	Sizes of disbursed loans < 500	500-5000	> 5000	Total
1992	573,631	1,734,374	169,164	2,477,169
	(23.16%)	(70.01%)	(6.83%)	(100.00%)
1993	1,204,141	2,814,441	249,212	4,267,793
	(28.21%)	(65.95%)	(5.84%)	(100.00%)
1994	2,874,960	5,170,278	501,757	8,546,995
	(33.64%)	(60.49%)	(5.87%)	(100.00%)
1995	4,916,332	8,255,053	1,526,004	14,697,389
	(33.45%)	(56.17%)	(10.38%)	(100.00%)
1996	5,919,510	12,840,243	2,705,629	21,465,381
	(27.58%)	(59.82%)	(12.60%)	(100.00%)
1997	5,345,655	17,901,772	5,224,518	28,471,945
	(18.78%)	(62.88%)	(18.35%)	(100.00%)
1998	5,430,785	22,402,337	7,323,485	35,156,607
	(15.45%)	(63.72%)	(20.83%)	(100.00%)
1999	4,782,291	23,755,821	8,207,702	36,745,815
	(13.01%)	(64.65%)	(22.34%)	(100.00%)
June 2000	2,401,577	13,445,719	5,188,436	21,035,732
	(11.42%)	(63.92%)	(24.66%)	(100.00%)

Table 3.27: Development of total amount disbursed by loan size over time.

Loan Size	Currency Boliviano	US Dollar	Inflation adj. Bolivianos
< 500	85.73	8.24	6.03
500-5000	29.16	65.65	5.19
> 5000	0.30	99.57	0.12

Table 3.28: Distributions of currencies by loan size (in %).

	Business Sector					
Year	Agriculture	Commerce	Stockbreeding	Production	Service	Total
1992	77	75		88	76	80
1993	63	95		100	87	96
1994	78	109		117	103	111
1995	121	145	141	158	133	147
1996	191	202	235	212	193	204
1997	250	290	265	300	305	293
1998	303	359	319	373	392	363
1999	360	455	325	466	502	458
June 2000	425	522	434	525	601	531

Table 3.29: Mean length of loans in days by year and business sector.

first, second ..	Year of first loan (cohort)						
approved loan	1992	1993	1994	1995	1996	1997	Total
1st	1,518	2,844	6,647	9,495	12,31	10,658	43,472
	(100)	(100)	(100)	(100)	(100)	(100)	(100)
2nd	1,333	2,429	5,6	7,697	9,47	7,702	34,231
	(87.81)	(85.41)	(84.25)	(81.06)	(76.93)	(72.26)	(78.74)
3rd	1,199	2,042	4,779	6,386	7,378	5,076	26,86
	(78.99)	(71.8)	(71.9)	(67.26)	(59.94)	(47.63)	(61.79)
4rth	1,071	1,723	3,997	5,193	5,255	2,609	19,848
	(70.55)	(60.58)	(60.13)	(54.69)	(42.69)	(24.48)	(45.66)

Table 3.30: Number of clients with a first, second, third, and fourth loan by year of first loan (%).

	Loan size				
Year	$<$100	100-300	300-1,000	1,000-10,000	$\geq 10,000$
1996	0.33	0.60	0.58	0.38	
1997	1.34	1.44	1.50	1.29	
1998	1.54	1.96	2.35	2.69	0.85
1999	2.40	2.89	3.63	4.13	4.46
2000	3.16	3.76	5.19	6.00	5.47

Table 3.31: Fraction of clients with arrears of $\geq$ 30 days by loan size (in %).

Chapter 4

The Impact of Microfinance Loans on the Clients' Enterprises: Evidence from Caja Los Andes, Bolivia

4.1 Introduction

Most developing countries have a large informal sector, constituted of small unregistered businesses. Since jobs in the official sector are scarce, people become micro-entrepreneurs, selling goods on the streets, adding to their income through home-production, or farming whatever piece of land available. The prevalence of the informal sector is especially large in South America, where many countries' legal system makes a formal registration extremely difficult. A 1983 household survey in Bolivia's capital La Paz found that 57% of the labor force was involved in the informal sector and that 89% of all retailers were informal (Rhyne 2001, p. 43). Wages and incomes in the informal sector are low and a large part of micro-enterprise workers live in poverty.[1]

Most micro-enterprises suffer from an inadequately low level of capital since their owners usually do not have access to the formal banking sector. Even if loans from the formal sector are available, transaction costs tend to be very high, estimated at 2% to 30% of the loan size (Murinde 1996, table 3.10). To alleviate short term

[1] An assessment of the micro-enterprise sector in South American can be found in Orlando and Pollack (2000). An account of the informal sector in Peru is provided in De Soto (1989).

capital needs, many of these households borrow from moneylenders who charge very high interest rates. In a survey of rural moneylenders in Nigeria, Aleem (1993) finds that the average annual rate charged was 79%. An evaluation report from a rural project in Vietnam finds that moneylenders unchallenged by credit programs charged around 15% per month (Kervyn 2001). The poorest households thus have extremely high borrowing costs.[2]

Politicians have attempted to reduce the interest rates paid by the poor through various measures. Some countries imposed interest ceilings, others set up development banks making available inexpensive loans. A large part of this money, however, never reached the targeted population. Since the funds were offered with highly attractive interest rates, a large proportion got diverted to more influential groups of the population. If the loans got distributed to the poor, they often were understood as gifts rather than loans and rarely were repaid.[3]

After the poor success of these development banks new approaches were tried. It was recognized that interest rates had to be raised to ensure an operation on a cost-covering basis and to make the loans unattractive to richer borrowers. While the rates have increased substantially, they are still considerably lower than the moneylenders'. In addition, new incentive schemes have been designed to generate high repayment rates, a prerequisite for long-term credit services. One of the first and most well known examples for these new institutions, the Grameen bank, was set up in Bangladesh in 1976 and serves over two million clients by now. Since most households cannot provide collateral, the Grameen bank distributes group loans only. All participants in the group are jointly liable for the loan amount and are granted consecutive loans only if the loan is repaid in full. The loan amount is increased over time, raising the potential gains from each consecutive loan. Through this mechanism the bank trusts in the group members to ensure that only reliable persons participate in the group and that the loan is repaid in full. The Grameen and other microfinance banks have achieved astonishingly high repayment rates through

[2]A slightly different view is expressed in Kochar (1997). Using data from a field study in rural India the author finds that borrowing constraints are less severe than commonly assumed.

[3]A comprehensive overview of the experience with development banks is given in Krahnen and Schmidt (1994).

these and related methods, becoming a model for development finance worldwide.[4]

At the time of writing this paper, there was an ongoing discussion among microfinance practitioners and supporters about the use of subsidies (see the articles and opinions in various issues of The MicroBanking Bulletin for some examples). One group argues that microfinance provides an excellent means to help the poor, who benefit from lower interest rates and improved access to loans. The costs of these small loans are very high and it is argued that the poor need help to cover these costs since they would have to pay unacceptably high interest rates otherwise. Others argue that microfinance can provide long-term services only if it eventually operates on a sustainable basis. While proponents of the first view want to push loan sizes further downwards to serve an even poorer clientele, the others argue that very poor households cannot be served on a cost-covering basis and frequently require aid rather than loans.

At the heart of this discussion lies the question how much clients actually benefit from microfinance loans and whether or not poorer households benefit more than others. As long as the majority of microfinance institutions uses subsidies in one form or another,[5] one has to compare costs and benefits of supporting microfinance programs to alternative ways of development aid. As long as the impact of microfinance programs has not been assessed, the discussion about the use of subsidies has to remain inconclusive.

How can the impact of these programs be measured? To begin with, microfinance institutions can offer substantially lower rates than moneylenders due to their larger scale of operations and lower costs of funds. The direct impact then can be measured by the reduction in borrowing costs. The biggest obstacle for more comprehensive studies analyzing the impact of the loans on the clients and their businesses is the lack of adequate data. Most institutions collect very little data from their

[4]For more information see The MicroBanking Bulletin, which regularly publishes key indicators for microfinance institutions worldwide. Morduch (2000) provides a recent overview of microfinance and discusses the underlying concepts. Since the emergence of group loans there have been a number of theoretical studies on the so-called "social collateral" analyzing the incentive mechanisms inside these groups, see Conning (2000) for an example. Practices in individual microfinance lending are discussed in Armendáriz de Aghion and Morduch (2000) and Churchill (1999).

[5]The MicroBanking Bulletin covered 49 institutions in October 1997 out of which 21 were fully self-sufficient. In September 2000, it covered 179 out of which 65 were fully self-sufficient.

clients. Since virtually none of the clients keep accounting books, information on their businesses is hard to come by. Even if this basic data is available, data on control groups almost never is. That is, even if we do observe the increase in income of the microlender's clients, we know nothing about non-clients and, in particular, do not know if and by how much their income increases as well. When we observe an increase in the clients' incomes, it is unclear how much of this increase is due to the loans and how much is due to the selection on the side of the bank. If the bank approves loan applications for relatively productive people, increasing incomes of clients might be induced by their higher productivity, independently of any loans. Finally, the placement of microfinance programs is endogenous. A microfinance institution opens its doors at a place where it is likely to achieve a raise in incomes. From increases in income in this location it is hard to infer whether a similar raise had been achieved in other places as well. This selection is particularly severe in countries where only few and relatively small microfinance institutions operate.

The objective of this paper is to analyze the impact of micro-loans on the clients' businesses. In particular, we analyze their contribution to productivity and growth. For the majority of clients, we find a positive impact of prior loans on productivity that increases with business size. That is, in terms of productivity, larger businesses tend to benefit more from microfinance loans than smaller businesses. Besides the analysis of productivity, the examination of growth is of particular relevance for an impact analysis. Many microfinance programs have been developed with the specific goal of increasing the scale of the clients' businesses. Many practitioners, on the other hand, suspect that the loans and the saving in borrowing costs are merely used for additional consumption instead of investment. If this were the case, neither assets nor incomes would rise. Our results, however, show a positive and significant influence of prior loans on growth in assets.

Throughout our analysis we address selection issues by estimating a two-stage selection model using clients with rejected loan applications as a control group. The analysis is based on data from Caja Los Andes in Bolivia. The data set includes valuable information about the clients' businesses in form of estimated balance sheets. Because of the widespread availability of micro-loans in Bolivia, the selectivity of the data is limited. Bolivia is one of the most developed and most competitive microfinance markets in existence today, Rhyne (2001, p. 19) estimates that between one third and one fourth of all Bolivian micro-enterprises are active borrowers.

The paper proceeds with an overview of related studies in section 4.2. Section 4.3 provides a brief description of the data set used and section 4.4 discusses the measurement of impact and the underlying theoretical concepts. Section 4.5 continues with a discussion of the econometric issues arising from the selection processes inherent in our data. Section 4.6 provides estimates for the selection equations and analyzes the impact of microfinance loans on growth and productivity. A discussion of the results is presented in chapter 4.7.

4.2 Literature

Due to a lack of suitable data, there exists only a small number of studies that analyze the impact of microfinance loans. There is a detailed micro-survey of households in Taiwan, including information on assets, loans, and savings. Besley and Levenson (1996) and Kan (2000) use this survey and find that a household's participation in an informal savings group (a rotating savings and credit association) has a positive impact on household investment.

In 1991/1992 the Bangladesh Institute for Development studies and the Worldbank jointly conducted a survey of households in rural Bangladesh to examine the impact of three microfinance programs. The survey was designed to include a number of control groups such as villages without access to one of these programs and households not eligible for participation. This survey has been used by a number of studies since. Pitt and Khandker (1998), for example, analyze the impact of group loans from one of these microfinance programs with a focus on gender-specific effects. They find that annual household expenditure increases by a larger amount if women are the recipients of these loans. In a follow-up analysis, Pitt, Khandker, Chowdhury, and Millimet (1998) examine the impact of group loans on children's health and find a significant positive impact of loans to mothers and a non-significant impact of loans to fathers. Morduch (1998), on the other hand, does not find any evidence for higher consumption levels or increased school enrollment. However, he does find evidence for lower variability of consumption implying that participating households better manage to smooth consumption over time. In an overview article (Morduch 2000) he discusses these and other findings. The effects of these programs on wages and employment are examined in Khandker, Samad, and Khan (1998) and Pitt (1999), who find evidence for increases in wages and self-employment. McKernan (2000) and Madajewicz (1999) analyze the impact of participation in these programs on profits. While McKernan finds a significant impact with profits increasing by roughly 175%, Madajewicz focuses on the distinction of group loans versus individual loans. She finds that when compared to individual loans, group loans from the Grameen bank increase profits by 8% for households with no land and by less for wealthier households (with a negative influence on profits for households with more than 2 acres of land). That is, wealthier households benefit more from individual loans than from group loans.

Coleman (1999, 2001) analyzes a microfinance program in Northeast Thailand. Correcting for selection bias, he finds that the impact of microfinance institutions on household wealth is either non-significant or negative. He attributes the negative impact to the small size of the loans. Being too small for investment, the loans are used for consumption and households turn to moneylenders to finance the repayments, leading to a vicious circle. When he distinguishes between wealthy and poor clients, he finds that only the wealthy clients benefit from the loans. The results by Coleman and Madajewicz have a similar structure in that they show the large influence of wealth. While the authors find negative or insignificant effects if averages are considered, there are significantly positive effects for groups with high wealth (Coleman (2001) and individual loans in Madajewicz (1999)) or low wealth (group loans in Madajewicz (1999)).[6]

In recent years, the "Assessing the Impact of Microenterprise Services" (AIMS) project has provided guidelines for impact analyses based on data already collected

[6] For further analyses of informal capital markets see Hoff, Braverman, and Stiglitz (1993), who provide a collection of articles on rural credit markets including many case studies, and Montiel, Agenor, and Ul Haque (1993) who provide an empirical and theoretical overview of informal financial markets.

by most institutions.[7] One strategy suggested consists in the comparison of clients who have passed the bank's screening but who have not yet received a loan, with clients who have received a loan some time ago. Through the restriction to clients who have passed the bank's screening, selection bias is reduced. Mosley (2001) and Copestake, Bhalotra, and Johnson (2001) use this approach to asses the impact of micro-loans in Bolivia and Zambia, respectively. Both find a positive impact of loans on the clients' economic situation and Mosley also finds evidence for poorer clients benefiting less because they prefer low-risk and low-return investments.

While these studies reduce selection bias in restricting their analysis to clients who have passed the bank's screening, their results are of a limited generality since they apply to clients only. That is, we cannot infer whether similar benefits would have been achieved if the program was extended to a larger part of the population. The benefits observed could be explained by the selection of the clients: if the bank selects those clients that make the best use of additional funds, we cannot expect similar benefits for other potential clients. In addition, neither paper models drop-out of clients. When comparing experienced clients with new clients these experienced clients have continued taking loans through some time which makes them a selective sample. The analysis presented in the following sections avoids both sources of bias through the estimation of a selection model.

While most of the above mentioned studies provide evidence for a positive influence of micro-loans on household welfare, they do not explicitly model the micro-enterprises through which the increase in incomes is achieved. One exception is McKernan (2000), who estimates a reduced form profit equation. While she finds that participation in a microfinance group increases profits, the analysis is restricted to contemporaneous effects. That is, profits are higher while participating in the program. The data do not allow inferences about longer term effects such as growth of the businesses. In addition, most of these studies use data from Asian countries

[7] One such guideline is "Learning from Clients: Assessment Tools for Microfinance Practitioners." The AIMS project is implemented by Management Systems International, Washington D. C., in partnership with the Harvard Institute for International Development, the University of Missouri, and the Small Enterprise Education and Promotion Network. For more information see *http://www.mip.org/componen/aims.htm*.

where the samples consist of very poor households in a restricted rural economic environment. Microfinance in South America, in contrast, caters to a different group of clients. The loans distributed are considerably larger and are targeted to the better-off households among the poor. Interest rates charged are higher and more institutions work on a cost covering basis.[8] As a consequence, we can expect the structure of the micro-enterprises and the way income is generated to differ considerably.

Besides the impact studies discussed above, there is a large body of literature on small firm growth in developed countries, most of which follows the debate on Gibrat's law (stating that growth is independent of firm size). Evans (1987a, 1987b), for example, uses data on US manufacturing companies and finds a negative connection between firm size, firm age, and growth. McPherson (1996) uses survey data from four African countries and analyzes growth determinants for a small micro-enterprise sample finding a significant influence of business sectors, human capital, gender, and firm size on growth. This paper adds to this literature in analyzing growth determinants for a large micro-enterprise sample focusing on the influence of prior loans on growth.

4.3 Data

The data we use for our analysis has been provided by Caja Los Andes, Bolivia. It consists of information on 76.000 clients and 28.000 rejected loan applications and covers the time from Mai 1992 to June 2000. Caja Los Andes FFP S.A. is a registered savings and loan company with its main branch in La Paz, Bolivia.[9] In July 2000 (when the data was collected) it offered credits to small and micro-enterprises in rural and urban areas. It also offered savings accounts and fixed deposits. Both loans and deposits are either in Bolivianos or in $US, acknowledging the widespread

[8] Of the institutions covered in The MicroBanking Bulletin (2000), 77% of the South-American institutions were financially self-sufficient compared to 55% of the Asian institutions. While only 39% of the South-American institutions specifically target the lowest income group (average loan balance $US 250), 79% of the Asian institutions do so (average loan balance $US 83).

[9] FFP stands for Fondo Financiero Privado. The legal category FFP has been created as an institutional form for small banks in Bolivia (Rhyne 2001).

dollarization of the Bolivian economy. In December 1999, Caja Los Andes was serving 36,815 clients with outstanding loans amounting to $US 35.9 Mio. 54% of these loans were made to women. The high concentration of micro-enterprises in the commerce sector is mirrored in the distribution of the outstanding loans, 44% of which went to commerce, 21% to production, 12% to the service industry, and 12% to agriculture related businesses. Since the data for agricultural businesses differs from the rest of the data, we confine the following analysis to the commerce, production, and service sectors.[10]

Caja Los Andes does not give group loans but secures the loans through chattel items such as televisions or other household items. While these have little value for Caja Los Andes, the owner values them very highly and has a strong incentive to repay the loan. Besides chattel items, personal guarantees are used and larger loans can be secured by mortgages as well. When a new client applies for a loan, the loan officer records the application. He visits the client's business and estimates balance sheet data if there are no obvious reasons for a rejection of the loan (these could be the age of the client, less than one year of business experience, or a bad repayment record with other banks). The loan officer then suggests whether and for which amount this loan should be approved and a committee decides (more experienced loan officers decide by themselves). When the client later on applies for another loan, the loan officer visits again and makes an update of the balance information. Clients with a very good repayment performance eventually obtain an automatic credit line and balance information is collected irregularly.

While Caja Los Andes initially gave loans to micro-enterprises only (i.e. very small enterprises), the target group has broadened in recent years. Caja Los Andes has opened a new small enterprise division which gives substantially larger loans. The median loan amount disbursed has increased from $US 367 in 1992 to $US 565 in January-June 2000. A part of this increase in loan sizes is due to a change in management policy: In spite of relatively high interest rates very small loans are

[10]The data for agricultural businesses contains information about assets and liabilities but does not contain information about income and expenditures. This is because a large part of the proceeds of agricultural businesses are consumed directly and the income generated is hard to measure.

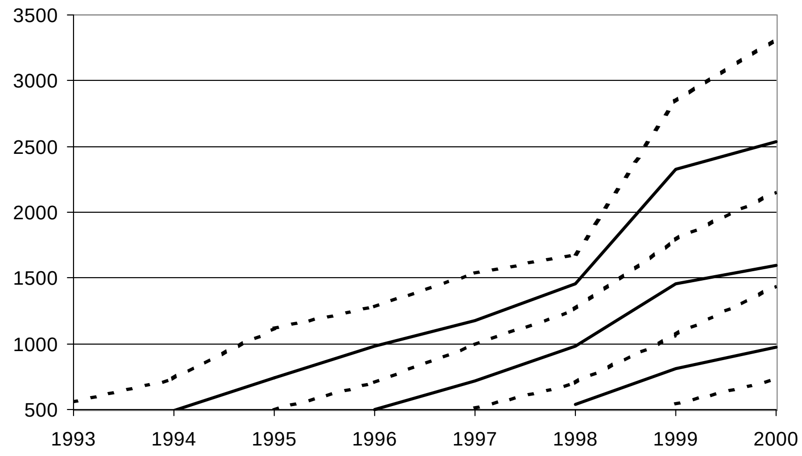

Figure 4.1: Median assets by cohort in 1992 $US. Sample: clients with initial assets below 1,000 (1992 $US), commerce, production, and service sectors.

hardly sustainable and larger loans are distributed to ensure cost coverage. The increase is also driven by an increasing need for funds from clients with fast growing businesses. The data show a substantial development of the clients over time as depicted in figure 4.1. Median assets for all clients who took their first loans in 1993 have increased from $US 600 in 1993 to $US 3,300 in 2000 (values are in 1992 $US).[11] While these numbers are impressive, it is unclear whether the increase can be attributed to the loans from Caja Los Andes. Clients might simply be selected based on their good growth prospects. Moreover, the data from 2000, for example, is based on those clients that take repeat loans in 2000. These are hardly representative for all clients who took a first loan in 1993. Section 4.5 presents the econometric theory underlying these selection processes and discusses a consistent estimator.

[11] The data is discussed in more detail in chapter 3. More information about the development of microfinance in Bolivia is provided in Rhyne (2001), Gonzalez-Vega, Meyer, Navajas, Schreiner, Rodriguez-Meza, and Monje (1996), Navajas, Schreiner, Meyer, Gonzalez-Vega, and Rogriguez-Meza (2000), and Inter-American Development Bank (1998). A lively account of Caja Los Andes' work can also be found in Frankfurter Rundschau (1998, 17. Oktober).

4.4 Measurement of Impact

Given the data, how can we measure the impact of loans from Caja Los Andes on the clients and their businesses? At the heart of this question lies the clients' use of their loans. While a part of these loans is invested directly into the clients' businesses and increases assets, many clients also use a part of their loan to finance consumption expenditures. When a client obtains a loan, he faces an intertemporal optimization problem for the decision how much of this loan to invest in his business and how much to consume directly. If clients were not capital constrained, the optimal level of assets would solely be driven by prices, expected returns, and interest rates, that is, it would be such that the marginal return on assets equals the interest rate paid on the loan. Since most of Caja Los Andes' clients do face constraints on the size of their loans, however, the levels of prior assets and of income (from their businesses and from other sources) play a crucial role for their decision how much to invest. To see this, consider an example where it would be optimal to use assets worth $1,000. If the client can freely chose the size of the loan, he will increase his assets to this level. If he obtains a loan of $200 only, for example, the size of initial assets and of his income determine whether it is optimal (and feasible) for him to increase assets to $1,000.

One possibility to measure the impact of the loan thus lies in an analysis of asset growth.[12] If the clients invest the additional funds and use them productively, assets should be higher than before even after the loan has been repaid. If it is optimal for the client to invest into his enterprise when he takes out a loan, it is optimal to use a part of the additional revenues for investment as well (assuming well-behaved preferences). If loans from Caja Los Andes contribute to an increase in assets, higher loans should lead to higher growth rates as long as the level of assets is sub-optimal (everything else equal). In addition, clients with a higher income from their businesses and from other sources are able to invest larger amounts than others.

[12] Alternatively, one could analyze growth in equity. Our information about liabilities, however, tend to be underestimated and we consider assets only. We do, however, subtract the amount of outstanding loans from Caja Los Andes.

To assess these factors, we estimate the determinants of growth in assets. In analogy to Evans (1987b) we estimate growth through a logarithmic expansion of log assets. Our estimation equation thus has the form

$$\frac{\ln K_{t'} - \ln K_t}{t' - t} = \ln A + \beta_1 \ln K_t + \beta_2 \ln^2 K_t + \beta_3 \ln^3 K_t, \tag{4.1}$$

where K are assets and A contains additional variables, among which are the number and log average size of prior loans and income. Since the time between two observations differs considerably, we use annualized growth rates. Assuming that the return on assets exceeds the interest rate, we expect that the size of prior loans has a positive influence on growth.[13]

Even if loans lead to an increase in the clients' assets, it is unclear whether or not the clients use these additional assets effectively. A second step of our impact analysis therefore examines the productivity of clients. In other words, how effectively do clients with prior loans use their additional assets? In a way, this analysis examines how smoothly clients can move upward on their production function. A priori, it is unclear to what extent the scale of a business can be increased smoothly. On the one hand, efficient use of larger assets might require other skills than running a small trade shop, for example. On the other hand, the bank's estimation of balance sheet information and the requirement of regular repayments might improve the clients' management skills and lead to higher productivity.

To examine the effects of prior loans on productivity, we compare new clients with an approved loan application but before the loan has been distributed with repeat clients which we observe after they have repaid their loans. Everything else equal, do repeat clients generate as many revenues from the additional funds as new clients do with their own funds? Given the same amount of assets, do repeat clients obtain higher or lower sales revenues than new clients?

We conduct the analysis through the estimation of a translog production function of the form

$$\ln(Y) = \delta D + \ln(A) + \beta_1 \ln(K) + \beta_2 \frac{\ln^2(K)}{2} + \beta_3 \ln(L) + \beta_4 \frac{\ln^2(L)}{2} + \beta_5 \ln(L)\ln(K) \quad , \tag{4.2}$$

[13]A formal analysis of the clients' intertemporal decision would reveal few additional insights relevant to the following analysis and is beyond the scope of this paper.

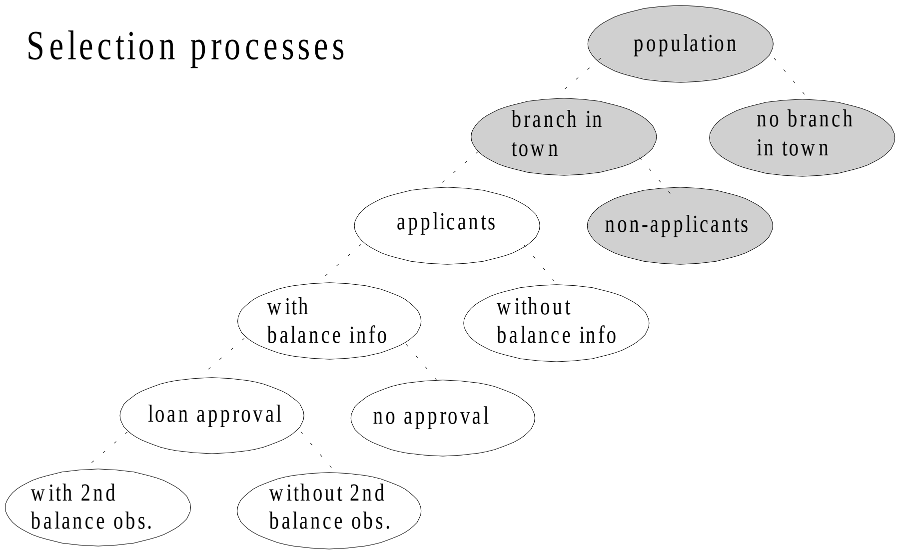

Figure 4.2: Step-wise selection. Shaded circles refer to unobserved groups.

where Y are sales, D is a dummy variable with a value of one if the client had prior loans, K are assets, L is the number of employees including the owner, and A is a measure of productivity determined by individual characteristics. The translog function is a generalization of the Cobb-Douglas production function, relaxing the assumption of a unitary elasticity of substitution, see Greene (2000, chapter 7.5). If clients use their additional assets efficiently, they should generate sales revenues which are at least as high as the new clients' (given the same asset level). δ determines the differences in sales revenues between new and repeat clients. $\delta \neq 0$ corresponds to a shift in the production function and the sign of δ shows whether or not the clients use the additional funds effectively.

4.5 Selection

The client data used for our analysis is no random sample of micro-enterprises. Not all micro-enterprise owners have access to a microfinance institution. Of these, only a part apply for a loan. While we can expect loan applicants to differ from the general population in a number of characteristics (location and economic environment, entrepreneurial skills, capital usage and willingness to incur debt, for example) we cannot incorporate these differences into our analysis since we do not observe the whole population. Clients enter our data set at the time of their application. After the submission of an application, the loan officer records additional balance information only if he takes a loan approval into consideration. Further along time, only successful clients can repay their loan on time and are granted consecutive loans. The longer the client stays with Caja Los Andes, the more selection processes he has passed. This step-wise selection can considerably bias the results of an impact analysis, it is illustrated in figure 4.2. The shaded circles refer to groups of the population we do not observe. In the following paragraphs we discuss the effects of two selection processes we observe (loan approval and the existence of a 2nd balance observation) in terms of their effects on enterprise growth.[14]

The first selection occurs when a client applies for a loan. There is self-selection on the side of the clients as well as selection from Caja Los Andes who approves only 73% of all loan applications. Quite possibly, clients whose enterprises have a high growth potential are more likely to obtain a loan than others. If this is the case, clients should have higher growth rates than non-clients independently of any loans from Caja Los Andes. If we contributed all the clients' growth to the loans, we would overestimate the effects of the loans. To incorporate this selection in our growth analysis, we can use data on rejected loan applications. Since there is no comparable database about the general population, we cannot model the clients' self selection.

A second source of selection bias lies in the existence of a second balance information which is recorded only if the client applies for a second loan. One could imagine that clients who had high growth rates in the past are more likely to apply for a second loan. As a consequence, the observation of high growth rates would not be driven by loans from Caja Los Andes but mainly by self selection on the side of the clients, leading to an over-estimation of the impact of prior loans on growth.

In the following paragraphs we formally describe the selection problem arising for the analysis of growth rates. Let $X = (X_0, X_1, X_2)$ denote the observed client

[14] A discussion of the collection of balance information follows in section 4.6.1.

characteristics. We are interested in a particular outcome y, growth in assets. y, however, is observed only if the client's first loan application is approved and if he applies for a second loan. Let z_1 denote a binary variable with a value of one if the loan is approved and let z_2 denote a binary variable with a value of one if a second balance observation is recorded. Formally, we are looking for

$$\begin{aligned} E(y|X) &= E(y|X,\ z_1 \cdot z_2 = 1) \ \cdot\ P(z_1 \cdot z_2 = 1|X) \\ &\quad + E(y|X,\ z_1 \cdot z_2 = 0) \ \cdot\ P(z_1 \cdot z_2 = 0|X) \quad . \end{aligned} \tag{4.3}$$

While we observe $(y|X, z_1 \cdot z_2 = 1)$ and can estimate $P(z_1 \cdot z_2 = 1|X)$ and $P(z_1 \cdot z_2 = 0|X)$, we do not observe $(y|X,\ z_1 \cdot z_2 = 0)$. One possibility to address this selection problem is through a latent variable model in analogy to Heckman (1976). Assuming a linear relationship and letting y^*, z_1^* and z_2^* denote the unobserved latent variables, we can write

$$y^* = X_0\beta + u \quad , \quad \text{with} \quad (u|X_0) \sim N(0, \sigma_u) \quad , \tag{4.4}$$

$$z_i^* = X_i\gamma_i + v_i \quad , \quad \text{with} \quad (v_i|X_i) \sim N(0, 1) \quad , \qquad i = 1,2 \quad , \tag{4.5}$$

$$y = \begin{cases} y^* & \text{if } z_1 = 1 \ \wedge\ z_2 = 1 \\ \text{not observed} & \text{if } z_1 = 0 \ \vee\ z_2 = 0 \end{cases} \quad , \qquad \text{and} \tag{4.6}$$

$$z_i = \begin{cases} 1 & \text{if } z_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \quad , \qquad i = 1,2 \quad . \tag{4.7}$$

We want to estimate $E(y|z_1^* > 0, z_2^* > 0)$. Letting ρ_{0i} denote the coefficient of correlation between u and v_i, this term can be modified to

$$E(y|z_1^* > 0, z_2^* > 0) = X_0\beta + E(u|v_1 > -X_1\gamma_1, v_2 > -X_2\gamma_2) \tag{4.8}$$

$$= X_0\beta + \sigma_u\rho_{01}\ E(v_1|v_1 > -X_1\gamma_1) + \sigma_u\rho_{02}\ E(v_2|v_2 > -X_2\gamma_2) \tag{4.9}$$

$$= X_0\beta + \sigma_u\rho_{01}\ \frac{\phi(X_1\gamma_1)}{\Phi(X_1\gamma_1)} + \sigma_u\rho_{02}\ \frac{\phi(X_2\gamma_2)}{\Phi(X_2\gamma_2)} \tag{4.10}$$

The step from (4.9) to (4.10) is based on the property $E(X|X > c) = \frac{\phi(c)}{1-\Phi(c)}$ for the mean of a truncated standard normally distributed variable, see Maddala (1983). For a more general discussion of two-stage selection models see also Tunali (1983).

The estimation then proceeds as follows. We first estimate (4.5) for $i = 1,2$ and use the linear predictions $X_1\hat{\gamma}_1$ and $X_2\hat{\gamma}_2$ to determine the inverse Mill's Ratios

$\hat{\lambda}_1 = \frac{\phi(X_1\hat{\gamma}_1)}{\Phi(X_1\hat{\gamma}_1)}$ and $\hat{\lambda}_2 = \frac{\phi(X_2\hat{\gamma}_2)}{\Phi(X_2\hat{\gamma}_2)}$. Equation (4.10) then can be estimated by a simple OLS regression with the additional terms $\hat{\lambda}_1$ and $\hat{\lambda}_2$.[15] With this stepwise estimation, we can consistently estimate β, $\sigma_{01} = \sigma_u\rho_{01}$ and $\sigma_{02} = \sigma_u\rho_{02}$. $\sigma_{01} > 0$ implies that a higher probability of the approval of the first loan application goes along with a higher expected growth rate y. $\sigma_{02} > 0$ implies that a higher probability of having a second balance observation goes along with a higher growth rate. If $\sigma_{01} = 0$ and $\sigma_{02} = 0$, selection bias is not a problem.

While the observation of growth rates is characterized by censored-sampling where we observe growth for a subset of clients only, the analysis of clients' sales revenues can use a richer set of data. In any given year, we observe balance information for new clients with approved loans and for clients applying for repeat loans.[16] Expected sales revenues for clients with no prior loans have to be conditioned on only one selection process (loan approval). At this moment, it is unclear whether the client will have a second balance observation at some later point in time or not. Expected sales for clients with one or more prior loans, however, have to be conditioned on two selection processes, loan approval and the existence of a second balance observation. Letting D denote a dummy variable with a value of one for second balance observations and zero otherwise, we can estimate the effects of prior loans analogously to (4.10).

$$E(y|D = 1 \vee D = 0) = X_0\beta + D \cdot \delta + \sigma_{u,v1}\ \lambda_1 + D \cdot \sigma_{u,v2}\ \lambda_2 \quad , \tag{4.11}$$

where δ measures the effects of prior loans. We can estimate equation (4.11) with the two-step estimator described above.[17]

[15] In a general model, correlation terms for v_1 and v_2 would also enter (4.9) and (4.10), see Maddala (1983) or Tunali (1983). For our data, however, we cannot reject the hypothesis of zero correlation between v_1 and v_2. That is, unobserved characteristics determining the probability of a loan approval are not correlated with the probability of the existence of a second balance observation.

[16] We also observe balance information for some rejected applications. We use this information to determine the selection process but exclude it from the production estimates to keep the model manageable.

[17] This model is restrictive in that it assumes that β_0 is identical for new clients and for clients after their first loans. Various alternative specifications, however, have shown little variation in the coefficients between those two groups. Our estimates include some interaction terms between D and X_0 to capture these differences.

This two-step estimation procedure has frequently been criticized because of the underlying normality assumptions and weak identification. Both lead to low robustness, that is, slight changes in specifications may alter the results substantially. In response to these issues, a number of semi- or non-parametric procedures have been developed in the last two decades. While these estimators avoid some of the problems of parametric estimation, there is growing evidence that the specification of the main equation is more relevant for the robustness of the results than the degree of parameterization. Parametric procedures like the one outlined above tend to perform well if the mean of the model is correctly specified, see Vella (1998) for a more comprehensive discussion. While the parametric approach does have some shortcomings, the data set used here includes many imprecisely estimated values. We choose the relatively simple approach outlined above despite its shortcomings since it allows us to control the results of the intermediate steps and to avoid spurious conclusions. In addition, we carefully design the estimation to ensure high robustness and identifiability.

While the two-step estimator is consistent, identification is weak if it is solely based on the functional form. This is because the inverse Mill's Ratio $\frac{\phi(X\gamma)}{\Phi(X\gamma)}$ is close to linear for a considerable range of values. In this case, one cannot distinguish between the selection process and the equation of interest. Technically, the correction term would be a linear combination of the other explaining variables, that is, the $(X'X)$ matrix does not have a full rank. The underlying problem is identical to the case of high multicollinearity: identification is weak and the variances of the estimated coefficients are high. One possibility to circumvent this problem is to impose exclusion restrictions. Through the inclusion of variables in the selection equation that are not included in the main equation, one ensures that the correction term is not a linear combination of the other explaining variables and that $(X'X)$ has a full rank.[18]

While the upper range of our estimates for $X\hat{\gamma}$ generally is beyond the linear range and the model would be identified by functional form alone, we use exclusion restrictions to improve the identification of the model. The choice of these variables

[18] See Tunali (1983) and Vella (1998) for a more detailed discussion of identification issues.

is crucial for the robustness of the results. If we exclude relevant explaining variables from the main equations, we will attribute an inordinately large part of their variation to the selection processes. Our data contain a few variables which can be used for exclusion restrictions, these are discussed in the following paragraphs.

When estimating the probability of loan approval, we include the amount applied for and a dummy variable for being on the black list. These variables are excluded from the estimation of growth in assets and from the estimation of sales revenues. The amount applied for is mostly driven by a lack of personal funds or by unexpected cash shortages, both of which do not determine the client's productivity. The black list is based on the client's credit history with other banks.[19] Repayment problems are frequently caused by unexpected expenses and strongly depend on the client's character. In addition, these entries frequently have been made some time before the application and thus do not determine the client's current business situation. The above arguments suggest that these variables are largely unrelated to current and future sales revenues and growth.

When estimating the probability of a second balance observation, we use the length of the prior loan in days, the client's highest number of days overdue, and the ratio of amount applied for over the amount approved as explaining variables. These variables are excluded from the estimation of sales revenues and growth in assets. The length of the loan is mainly determined by the clients' repayment capacity and by the use of the loan. The calculation of the repayment capacity is based on prior incomes only, that is, it does not take into account the additional income generated with the help of the new loan. The length of the loan should thus be unrelated to future sales and growth in assets. Repayment behavior is partly determined by character, and partly by unexpected changes in the clients' expenses and income. The ratio of amount applied for over the amount approved is driven by differences between the client and the loan officer in assessing the client's need for funds and his repayment capacity.

[19] Although being on the black list should imply an immediate rejection of the application, this is not always the case. 0.4% of all approved loan applications are for clients on the black list.

4.6 Estimation

This section begins with a description of the selection equations for the two stages described above. We then incorporate the predicted selection to calculate impact estimates. Finally, we discuss differences between micro-enterprises and larger enterprises.[20]

4.6.1 Selection Estimates

As discussed in section 4.4, we estimate the probability of being selected at each stage with a Probit model. The results are discussed in the following paragraphs.

4.6.1.1 Loan Approval

When estimating the probability of a loan approval we first have to decide on which data this estimation should be based. We have basic data such as the amount applied for and business sectors for all clients. In addition, there is balance information for all approved loan applications and for roughly 7% of all rejected loan applications. This leaves us with two possibilities to determine the structure of the selection bias:

a. To infer the selection process from the balance information available treating the 7% of rejected applicants as a random sample of all rejected applicants.

b. To infer the selection process from the basic data available for all applicants.

Both variants have shortcomings. Treating the 7% as a random sample as in (a) yields biased estimates since these 7% have passed a first selection: the loan officer takes a loan approval into consideration. If we use the basic data only, as in (b), the predictive power is very low and the results suffer from omitted variable bias since we

[20]For the estimation, we adjust the sample as follows. To reduce the impact of outliers we drop observations whose assets are above the 99th or below the 1st percentile. Loans and balance observations were kept only if the dates matche and the calculated "income after repayment" is consistent (leading to a loss of 25% of our observations). The analysis is restricted to the first and second balance observations of each client to include as few selection processes as possible. Balance observations are required to be at least 90 days apart, otherwise the more recent one is dropped. Estimates for the second selection process (existence of a 2nd balance observations) are calculated for those clients only who had their first loan in or before 1997 to allow sufficient time for a second balance observation to occur. The year 1992 is excluded.

cannot use important factors such as liabilities over assets or guarantee information as explaining variables. Keeping in mind the shortcomings of both approaches we will calculate two sets of estimates, one corresponding to each (a) and (b).[21]

The results of the probit estimates can be found in table 4.4. The estimated coefficients show that the influence of the amount applied for is highly non-linear, with highest values for loan amounts slightly above $US 100 and decreasing thereafter. Clients in the commerce sector are more likely to obtain a loan than clients in the production sector. Having a bad repayment record with other banks (being on a "black list", which is based on information from the banking supervisory authority in Bolivia) and being single have a significant negative impact on loan approval. When we base our estimates on observations with balance information (column (a) in table 4.4), we find that being single no longer has a significant negative impact. In addition, the probability of approval increases for a lower liabilities over assets ratio.

When comparing variable means between rejected applications and approved applications, we find that liabilities over assets are higher for rejected applications (0.06 compared to 0.03). Rejected applicants are on the black list more frequently (1.8% compared to 0.4%). They are single more often (71.9% compared to 19.7%) and there are fewer women (56.2% compared to 61.2%). They are less frequently in the commerce sector (50.3% compared to 54.0%) and more frequently in the service sector (29.7% compared to 17.5%). For more details see table 4.5.

4.6.1.2 Existence of a Second Balance Observation

In a second step, we estimate the probability of the existence of a second balance observation given that the first loan application has been approved (table 4.6). We find that a second balance observation is more likely for larger approved loans. It is less likely for a high ratio of the amount applied for over the approved amount. That is, if Caja Los Andes gives a loan that is considerably smaller than the amount

[21] One could think of a selection model with three selection stages. Since the decision to record a balance observation and to approve a loan are driven by the same considerations, however, the identifying power of such an approach would be very low.

asked for by the client, this client is less likely to have a second balance observation. There could be two possible reasons: either the client chooses to take (larger) loans from another source and does not want to stay with Caja Los Andes, or Caja Los Andes gives relatively small loans to clients who are more likely to have late payments and thus do not obtain future loans. In addition, the probability of a second balance observation is higher for women, for non-singles, and older businesses. These characteristics tend to go along with stability and are frequently used by loan officers to assess the credit-worthiness of clients. The probability of a second balance information diminishes strongly if the client has a bad repayment record. The higher the maximum number of days overdue, the lower the probability of a second balance information. This dependency reflects Caja Los Andes' policy of rejecting applications from clients with a bad repayment record.

When comparing variable means, we find the most pronounced difference in the maximum number of days in arrears. While the average is 33.5 days for clients without a second balance observation, it is 2.4 days for clients with a second balance observation. Among clients with a second balance observation there are more women (64% compared to 59%) and businesses tend to be older (1.62 compared to 1.56 in logarithms). For more details see table 4.7.

4.6.2 Impact Estimates

This section quantifies the effects of loans from Caja Los Andes on its clients' enterprises. Section 4.6.2.1 analyzes how the number and average size of prior loans contribute to the growth of the micro-enterprise and section 4.6.2.2 continues with an examination of production. Everything else equal, we ask whether clients with prior loans generate higher sales revenues than other clients.

4.6.2.1 The Impact of Microfinance Loans on Growth

This section analyzes the determinants of micro-enterprise growth and asks whether prior loans contribute to growth. The question of growth in assets is of particular relevance sine it constitutes a prerequisite for an increase in business income. In addition, when we can show positive growth effects of prior loans we can infer that the savings in borrowing costs and the additional funds are used productively and lead to a lasting expansion of the business beyond providing additional short term capital.

To examine whether microfinance loans contribute to business growth, we estimate annualized growth rates for assets between the clients' first and second balance observations as discussed in section 4.4. To capture the influence of prior loans we generate two variables indicating to what extent the client has used loans from Caja Los Andes. These variables are the number of approved loans from Caja Los Andes prior ($\geq$ 60 days) to the balance observation (NUMPRIAP) and the average size of approved prior loans (APRISIZE). To incorporate the two sources of selection bias into our estimates, we follow the steps described in section 4.5 and use the selection estimates discussed in section 4.6.1.

The results of the growth estimates corresponding to equation (4.1) are presented in table 4.8. There are considerable size effects with growth rates being higher for lower initial assets and growth being higher for clients with higher income. The log of the number of days passed between the two balance observations, $\ln(time)$, is highly significant with a negative coefficient. That is, average annualized growth tends to decrease if the time span between two balance observations is relatively large. It is unclear whether this is due to a strong immediate impact of a loan which diminishes over time or due to the endogenous decision when to apply for a new loan. Businesses with high growth rates in the past might need to apply for a new loan earlier than businesses with lower growth. In addition, loan officers might decide not to record a new balance observation if the business has undergone few changes.[22] Growth is between three to four percentage points lower for female clients and decreases with business age.

The number and average size of prior loans have a significant influence on growth rates. The effect of the number of prior loans varies by sectors and with assets. While it is insignificant in the service sector, its coefficient is negative in the commerce

[22] As a check for the robustness of our results with respect to endogeneity of $\ln(time)$, we ran control regressions without this variable and found no mentionable changes in other coefficients.

and production sectors with a positive interaction effect with assets. That is, the combined effect is negative for small businesses, positive otherwise. The average prior loan size (APRISIZE) has a positive coefficient which is significant in all sectors. Besides the absolute size of prior loans, the size of these loans relative to assets also has a significant influence. Taken together, these effects imply that clients with low initial assets benefit from larger loan sizes but not from taking multiple loans of the same size. As an example for the influence of loan sizes, figure 4.3 compares the influence of one prior loan of $US 600 to one prior loan of $US 500 (all values are in 1992 $US). Given two clients that are identical with respect to all measured characteristics, but differ in their loan sizes, the client obtaining the larger loan has between $200 and $400 in additional assets after one year.[23] The additional increase in assets is largest for clients with relatively low initial assets and in the commerce sector. While a stronger increase in assets for clients with larger loans is not surprising at first sight, this increase is calculated after the loans have been paid back. That is, the increase in assets has to be financed by additional income generated through prior loans.

The large effect of prior loans on growth has to be taken with a grain of salt, however. While our estimates include correction terms for the approval of loan applications, we do not correct for the size of the loans approved. The average size of prior loans is partly determined by the loan officer's assessment of the client's repayment capacity and partly by the client's perceived need for additional funds. While the former is unrelated to future growth rates (the repayment capacity is calculated without taking the additional income through the loan into account), the latter is higher for clients planning to expand their businesses. APRISIZE thus is partly endogenous and we have to interpret the effects cautiously. The number of prior loans could also be driven by the loan officer's assessment. There are multiple loans between two consecutive balance observations only if the loan officer did not record a new balance observation when the client applied for a second loan (or when the balance information has been deleted or overwritten, which has happened

[23]We consider clients with initial assets between $1,000 and $5000 only since a loan size of $500 is most frequently observed in this range. Outside of this range, additional assets are slightly higher.

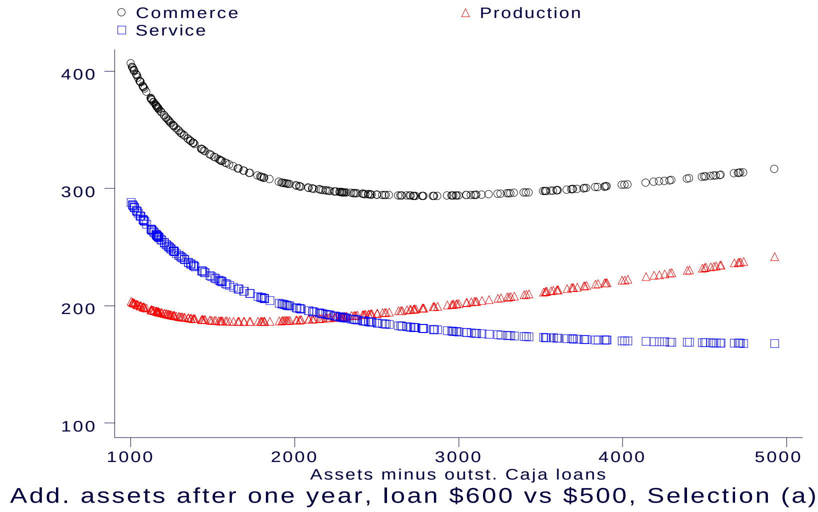

Figure 4.3: Additional assets after one year for a loan of $US 600 compared to a loan of $500. Calculations are based on columns 1, 3, and 5 in table 4.8. Sample: First loan in or before 1997 with a corresponding balance observation in or after 1993, first application not rejected.

frequently). The loan officer might not record a new balance observation if, for example, the business has not changed very much since the last visit. If this is the case, a high number of prior loans is a signal for few changes and the coefficients of NUMPRIAP are subject to endogeneity bias. Recognizing the potential bias of these variables, the production analysis in the next section uses a cross-sectional approach and thus avoids the use of these variables.

The effects of selection bias can bee seen from the coefficients for the two correction terms. The coefficient of λ_1 is mostly insignificant (positive and significant in the production sector) and the coefficient of λ_2 is negative and significant at the 1% level in the commerce and production sectors. In other words, the estimated coefficients indicate that Caja Los Andes selects clients who tend to have high growth rates in the future while the probability of having a second balance observation is negatively correlated with growth. There are unobserved characteristics which at the same time make it more likely to observe a second balance observation and low growth rates. One possible reason for this effect could be that the most successful

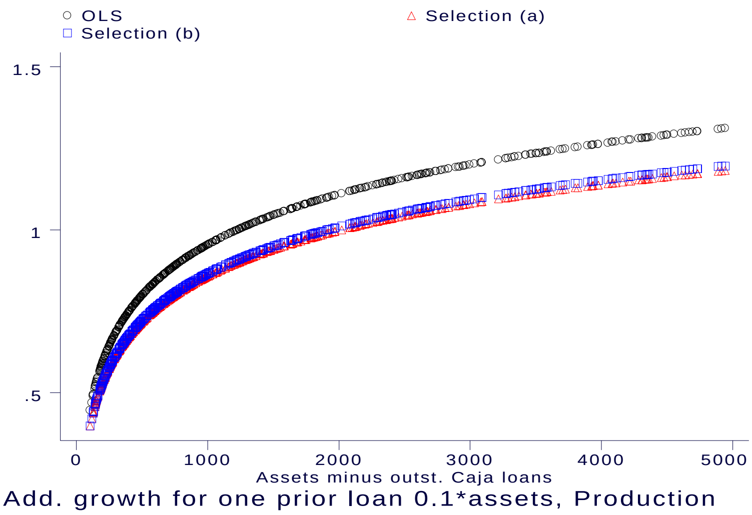

Figure 4.4: The effect of one prior loan on growth in assets in the production sector, when the loan equals 10 percent of assets. Sample: First loan in or before 1997 with a corresponding balance observation in or after 1993, first application not rejected.

clients need larger loans and look for other sources. The incorporation of the selection effects changes the estimated influence of prior loans as depicted in figure 4.4, where we compare the effects of one prior loan of a size corresponding to 10 percent of prior assets. The estimated effects on growth are highest for the OLS estimates, while a consideration of the selection effects reduces these effects.

One reason why we are interested in asset growth is that it is a prerequisite for growth in business income. To asses the extent to which growth in assets translates into growth in business income, table 4.1 shows the median of growth in business income relative to growth in assets by previous asset sizes. We find that the median ratio is close to 0.3. That is, if assets grow by 10 percent, business income grows by 3 percent. Businesses of a median size ($US 500-5000) are the best in translating growth in assets to growth in business income, although the variation between different asset ranges is low.

While growth in total assets is a good measure for the size of a commerce business, fixed assets might be more appropriate for the production sector since they determine the production capacity. In addition, asset growth could also be driven by unwanted

Asset range (1992 \$US)	<300	300-500	500-1,000	1,000-5,000	>5,000
median(growth in business inc. /growth in assets)	0.27	0.29	0.31	0.31	0.29

Table 4.1: Growth in business income relative to growth in assets for various asset ranges.

inventory buildup and does not necessary mirror an increase in the scale of the business. To acknowledge these sectoral difference, we report separate estimates for growth in fixed assets in the production sector in table 4.9. The main results remain valid, while the estimation has a lower explanatory power and levels of significance are reduced.

4.6.2.2 The Impact of Microfinance Loans on Productivity

This section examines micro-enterprise production. To compare the use of assets between new and experienced clients, we analyze sales revenues. Using the balance and client information available, we can estimate a translog production function according to equations (4.2) and (4.11) by the two-step procedure discussed in section 4.5. We report results separately for the commerce, service, and production sectors. Since the revenue structure differs considerably in the years of the economic crisis (1999 and 2000), we drop these years. We use the probit estimates discussed in section 4.6.1 to calculate $\hat{\lambda}_1$ and $\hat{\lambda}_2$. The productivity estimates are reported in tables 4.10 to 4.12.

Our model has a relatively high explanatory power for the commerce and production sectors with R^2 values of 63% and 55%, respectively. The processes in the service sector, however, are not explained as well by our model with an R^2 of 31% only. The explanatory power suffers from incomplete information on labor use where we observe only the number of employees and not the hours worked. While a relatively large part of sales in the service sector is determined by the number of employees, the amount of assets is more relevant for production and commerce sectors. When testing the translog specification we can reject the hypothesis that $\beta_2 = \beta_4 = \beta_5$ at the 1% level.[24]

[24] Among the micro-enterprises in our sample, 96% have no employees. When we restrict the

	Selection (a)	Selection (b)	OLS
Commerce	3.7	4.6	1.4
	*(3.74)***	*(4.55)***	*(2.09)**
Production	11.9	12.7	5.9
	*(8.27)***	*(8.80)***	*(6.46)***
Service	15.2	16.1	8.9
	*(6.82)***	*(7.23)***	*(6.15)***

Table 4.2: Average percentage increase of sales revenues for clients with prior loans by sector and estimation method. The number corresponds to the estimated coefficient on D from the first three columns in tables 4.10, 4.11 and 4.12. Robust t-statistics are in parentheses. * indicates significant at 5%; ** significant at 1%. For more details see the tables in the appendix.

Given the same amount of assets and employees, which type of client generates higher sales revenues? The results displayed in tables 4.10 to 4.12 show that sales revenues are higher for older clients, for women in the service sector, men in the production sector, and for older businesses in the commerce and production sectors. The positive effects decline with asset size. We also find a significant positive influence of liabilities over assets. Everything else equal, clients with higher liabilities generate higher sales revenues. This positive connection suggests that clients who take loans tend to be more productive than other clients. The results also show that there is a positive correlation between the decision of a loan approval (λ_1) and sales revenues while there is a negative correlation between the existence of a second balance observation (λ_2) and sales. That is, clients who have a second balance observation tend to be less productive than other clients. Both correction terms are significant at the 1% level.

We now turn to our main question: how effectively do clients use their additional funds? Given the same amount of assets and employees and all other characteristics being identical, do clients with prior loans obtain sales revenues comparable to those of new clients? This difference in sales revenues goes beyond an increase in

sample to these businesses, the main results remain unchanged. An additional check for robustness using other functional forms revealed no changes in the main results.

assets made possible through the loan. It is solely driven by changes in productivity e. g. through better accounting practices made necessary for the repayment of the loan. We can measure the influence of additional prior loans through the coefficient on the dummy variable for observations with prior loans (D). In a first set of regressions we restricted the influence of prior loans to a proportional increase in sales revenues. Table 4.2 lists the estimated coefficients for the three sets of estimates and for each sector. Taking into account the selection effects (columns 1 and 2) we find that commerce businesses with prior loans have approximately 4% higher sales revenues than businesses without prior loans, production businesses have 12% higher sales revenues, and service businesses have 16% higher sales revenues. All coefficients are significant at the 1% level. When we compare these results with simple OLS estimates, we find that the OLS estimates underestimate the impact of prior loans. While this underestimation is surprising at first sight, it corresponds to the estimated selection effects. If the clients that we observe after they have taken out a first loan are on average less productive than others, a simple comparison of these clients with new clients must underestimate the effects of the loans. More details can be found in tables 4.10, 4.11 and 4.12.

To take a closer look at the influence of prior loans on different types of clients, we consider an interaction with log assets and with the log of the number of days as a client. The coefficients are significant at the 1% level for the commerce sector and partly significant for the production and service sectors. While there is a positive influence of assets, the time a client has been with Caja Los Andes has a negative influence. That is, given the same size of assets, the positive impact on sales revenues is largest for clients with a relatively short time between their first and second balance observation.[25] Figure 4.5 depicts the combined effects of these variables for a client who has been with Caja Los Andes for 240 days. The effects of prior loans are negative for small businesses in the commerce sector, mostly positive otherwise and increase with assets. Clients with prior loans thus have experienced a shift in

[25] As mentioned above, the coefficient of time has to be interpreted cautiously. Since the time passed between two balance observations is driven by the length of the prior loan and the loan officers' decision when to record another balance observation, there could be endogeneity bias. Dropping $\ln(time)$ from the regressions we find that the other coefficients remain similar.

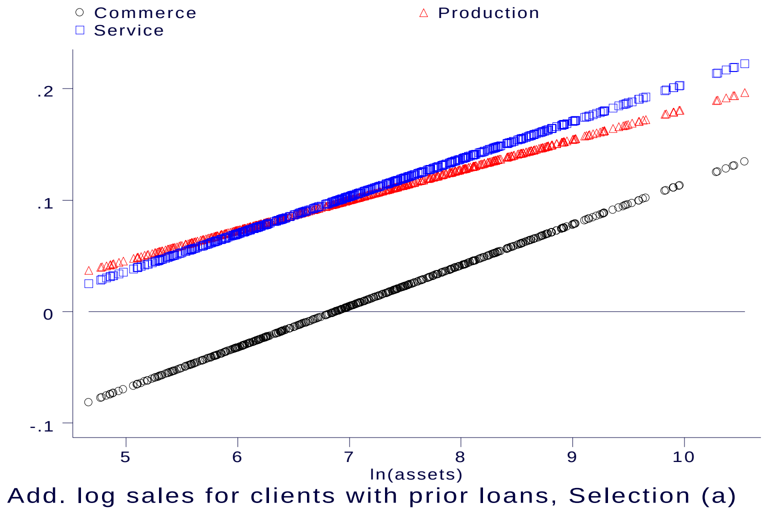

Figure 4.5: The influence of prior loans on log sales by log assets and sector. Sample: balance observation in 1993 to 1998, first application not rejected. 2nd balance observations are included only if the first balance observations was in or before 1997. The graphs are based on the coefficients of D, $D \cdot ln(assets)$, and $D \cdot ln(time)$ from the fourth columns of tables 4.10, 4.11 and 4.12, assuming $time = 240$.

their production function. That is, given the same amount of assets and all other measured characteristics being identical, clients with prior loans generate higher sales revenues than others. Figure 4.6 depicts the shift in the production function for the production sector.

From our estimated production function, we can determine marginal sales revenues for new clients without employees as

$$\frac{\partial Y}{\partial K} = A \cdot \frac{\beta_1 + \beta_2 \ln(K)}{K} \cdot K^{\beta_1 + \beta_2 \frac{\ln(K)}{2}} \cdot e^{\epsilon} = \frac{\beta_1 + \beta_2 \ln(K)}{K} \cdot Y \quad , \tag{4.12}$$

where β_1 has to be adjusted for interaction terms. From our parameter estimates it follows that marginal sales revenues are decreasing for the asset range considered. We can calculate the predicted marginal revenues for an example. Consider a male new client in La Paz in the first quarter of 1996 aged 30 with a business aged 4 years. Figure 4.7 shows the predicted monthly marginal sales revenues by asset size and business sector.[26] It is as high as 1.50 for small businesses in the commerce

[26] For the shape of the production function for all three sectors see figure 4.8 in the appendix.

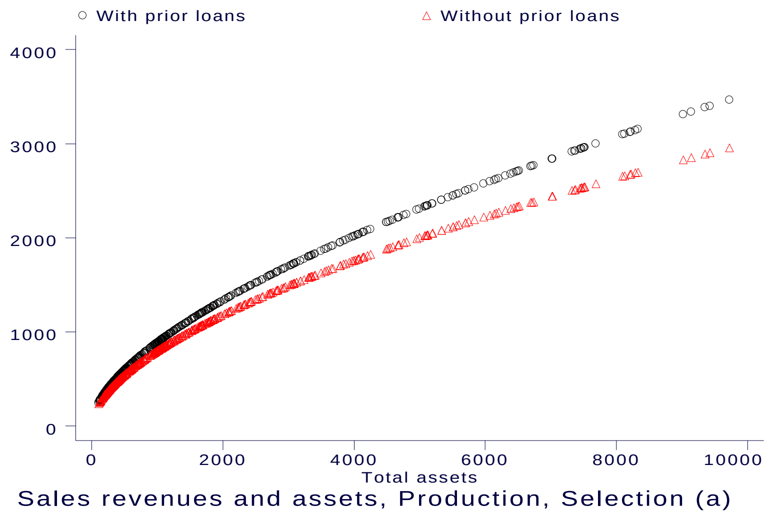

Figure 4.6: Estimated link between sales and assets for production businesses with and without prior loans. Sample: balance observation in 1993 to 1998, first application not rejected. 2nd balance observations are included only if the first balance observations was in or before 1997. The graphs are based on the fourth column of table 4.11, assuming $time = 240$.

sector. It is lowest for the service sector with values below 0.2. That is, out of 100 additional dollars in assets, clients in the commerce sector generate up to 150 dollars additional sales revenues per month. In other words, these clients turn over any additional inventory in less than a month.

In sum, we find a significant influence of prior loans on sales revenues in the commerce and production sectors. Prior loans increase sales for sufficiently large businesses. The puzzling result that smaller commerce businesses might not benefit from prior loans calls for a further analysis.[27]

[27]While the above analysis has been confined to sales revenues, the results have direct implications for business income due to the duality of cost functions and production functions. We estimate production functions rather than cost functions since a thorough estimation of costs functions would require disaggregated price data which is not available for our data set. Robustness checks have confirmed that the influence of prior loans on costs (and, thus, profits) has a similar structure as the influence on sales.

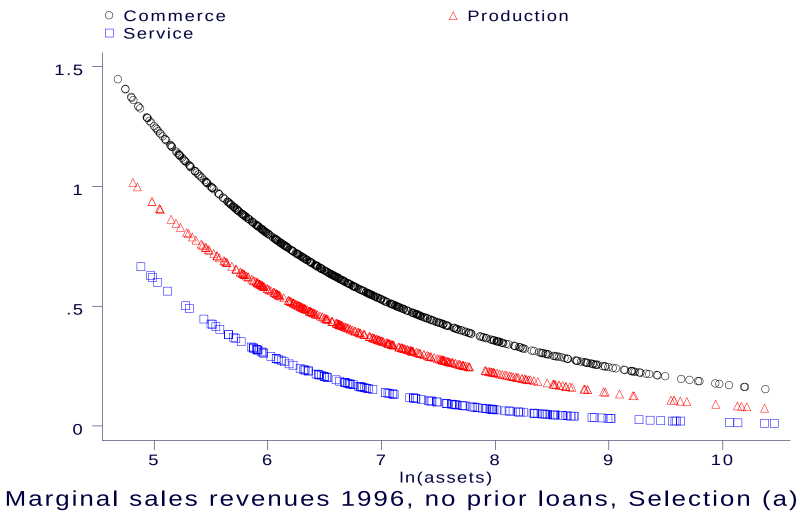

Figure 4.7: Predicted monthly marginal sales revenues by business sector for a male client in La Paz in 1996 with 30 years of age and a 4 year old business.

4.6.3 Size Effects

The results above have shown a number of differences between very small and relatively large enterprises. While there are pure size effects with growth rates as well as marginal sales revenues being considerably higher for smaller enterprises, the analysis of the clients' production structure showed a stronger impact of prior loans for larger businesses. While larger businesses can achieve sales revenues by up to 20% higher than similar businesses without prior loans, smaller businesses especially in the commerce sector do not use the additional funds as effectively. Compared with similar businesses without prior loans, they obtain up to 8% lower sales revenues.[28]

How can we explain this difference? Clients with small businesses might not be able to use the additional assets as effectively as other clients. While one can increase the scale of any given business to some extent, a larger increase often requires a new location, new techniques etc. which might not be feasible for very small businesses, whose owners lack the experience for an operation at a larger scale. This

[28]See the coefficients of D, $D \cdot \ln(assets)$, and $D \cdot \ln(time)$ in columns 4 and 5 of table 4.10, with $time = 240$ the effects are negative for businesses with assets below 900$US (column 4) or 350$US (column 5) in 1992 values. When we split the sample into different sub-groups by asset size the effects remain similar.

could be especially true in the commerce sector where the clients might prefer to sale at the same spot with the same neighbors even when they could earn more by moving to a larger location. This could also explain the observation that growth in business income relative to growth in assets is lower for small businesses, see table 4.1. Another explanation for the lower sales revenues of small clients could lie in the selection of clients. Our selection estimates show that clients applying for small loans are more likely to obtain a loan than other applicants (see the coefficients on the amount applied for in columns (a) and (b) of table 4.4). If these small businesses were chosen because they face the highest credit constraints and not because they have good repayment capabilities, they might not use the additional capital as effectively as larger businesses who were selected based on their repayment capability alone. A possible benefit of the access to loans could consist in improved accounting practices. Through the regular visits of the loan officers and the regularity of required repayments, clients might be forced to improve their control of costs and expenses. This effect most likely is stronger for larger businesses.

While small businesses do not benefit as much in terms of additional sales revenues than larger ones, they benefit in other ways. The calculation of marginal sales revenues has shown that the monthly turnover from assets to sales revenues is as high as 150 percent for small commerce businesses. That is, even if the small businesses might not use the additional assets very effectively when compared to other small businesses, these assets are converted to sales revenues at a much faster rate than in larger businesses.

4.7 Conclusion

The above analysis has examined the influence of loans from a Bolivian microlender, Caja Los Andes, on the clients' enterprises. We have shown that these loans raise productivity and growth for most clients. In particular, we find a significant permanent impact of prior loans on assets. That is, assets remain higher than before even after the repayment of the loan with growth rates in assets being higher for larger loans. The analysis of the production structure has shown that clients with prior loans from Caja Los Andes tend to have higher sales revenues than clients with no prior loans, where the difference increases in asset size. Clients in the production sector with at least one prior loan from Caja Los Andes generate up to 20 percent higher sales revenues than new clients with the same amount of assets and employees. In the commerce sector, clients with prior loans and sufficiently large initial assets generate up to 12 percent higher sales revenues than new clients. Smaller commerce businesses, however, seem not to use the additional funds as effectively. We found that clients with small enterprises in the commerce sector and one or more prior loans generate up to 8 percent lower sales than clients without prior loans and the same amount of assets and employees. While these clients are not very efficient in the use of their assets when compared to other small businesses, we also found that they have a very high turnover rate. That is, additional assets are turned into sales revenues at a much faster rate than in larger businesses.

Our estimators explicitly incorporate selection effects. While we found a positive correlation between loan approval and growth and productivity, we surprisingly found a negative correlation between the existence of second balance observations and growth and productivity. That is, the "best clients" discontinue borrowing after the first loan more often than others. Since an increasing number of impact studies resorts to "before-after" comparisons and thus relies on the sample of repeat borrowers (Mosley 2001, Copestake, Bhalotra, and Johnson 2001, for example), the link between productivity and continuous borrowing from the same lender should be further explored.

While the estimators were designed to correct for the biases arising from the bank's decision to approve a loan application and from the client's decision to apply for repeat loans, we cannot correct for selection bias arising from the client's decision to apply for a loan since we observe applicants only. When interpreting our results we thus have to restrict the analysis to micro-enterprises willing to take loans. For a more comprehensive impact study it would be desirable to have information on a randomized control group of rejected applicants and of the general population.

Our results have shown evidence for a positive impact of micro-loans on the enterprises. However, there is no evidence yet on market wide impacts such as changes in wages or changes in the incomes of non-participants, it is unclear whether micro-enterprises eventually provide income above subsistence levels, and there is very little information about the role of micro-enterprises for economy-wide growth. Further research could continue in two directions. Firstly, one could examine the merits of supporting micro-enterprises in comparison to larger and potentially more efficient enterprises. To what extent is it advisable to support micro-enterprises when larger enterprises are more efficient? Secondly, it would be interesting to extend the impact analysis from the micro-entrepreneur's perspective to a general equilibrium framework where market-wide changes are considered.

4.8 Appendix

Personal Characteristics	
D(single)	Marital status = single
D(female)	Gender = female
D(on black list)	Bad credit record with other banks
ln(age)	Log of the client's age
ln(non-business inc.)	Log of non-business income
ln(business income)	Log of business income
Previous maximum arrears	Maximal arrears in previous loan
NUMPRIAP	Number of prior loans
APRISIZE	Average size of prior loans
ln(time)	Log of the number of days since the first loan
D	Dummy: experienced client with ≥ 1 prior loan
Business Characteristics	
ln(assets)	Log assets
ln(employees)	Log number of employees including owner
Liabilities/assets	Ratio of liabilities over assets
ln(Business Age)	Log business age
D(Commerce)	Dummy: Commerce Sector (relative to production sector)
D(Service)	Dummy: Service Sector (relative to production sector)
Loan Characteristics	
ln(approved amount)	Log approved amount (loan size)
Appl./appr. amount	Amount applied for over approved amount
Length of loan (days)	Duration of the loan in days
ln(value of chattel g.)	Log value of chattel guarantees
Environment	
D(Cochabamba)	Dummy: loan disbursed in Cochabamba (relative to La Paz)
D(Sucre)	Dummy: loan disbursed in Sucre (relative to La Paz)
D(Santa Cruz)	Dummy: loan disbursed in Santa Cruz (relative to La Paz)
D(Trinidad)	Dummy: loan disbursed in Trinidad (relative to La Paz)
D(Tarija)	Dummy: loan disbursed in Tarija (relative to La Paz)
GROWTH	Quarterly growth rate (source: INE)
D(199x)	Dummy: Year=199x

Table 4.3: List of variables used for the empirical analysis. All logs are calculated as log(<variable>+1).

	P(Loan approval)			
	(a)	(b)	(a, $\leq$1,000)	(a, >1,000)
ln(amount applied for)	1.890	2.634	-0.415	2.552
	(2.83)**	(13.61)**	(0.28)	(2.43)*
ln(amount applied for)^2	-0.257	-0.380	0.194	-0.396
	(2.33)*	(11.86)**	(0.73)	(2.39)*
ln(amount applied for)^3	0.009	0.016	-0.020	0.017
	(1.50)	(9.07)**	(1.26)	(2.04)*
liabilities/assets	-0.327		-0.172	-0.512
	(2.55)*		(1.19)	(3.39)**
ln(business age+1)	-0.029		-0.029	-0.043
	(1.39)		(0.90)	(1.52)
D(single)	-0.016	-1.353	-0.011	-0.027
	(0.38)	(113.40)**	(0.16)	(0.51)
D(female)	-0.013		0.029	-0.021
	(0.37)		(0.51)	(0.47)
D(on black list)	-1.349	-0.887	-1.333	-1.382
	(11.33)**	(13.63)**	(7.37)**	(8.31)**
D(1993)	1.839	2.857	2.012	1.455
	(12.89)**	(24.78)**	(9.31)**	(7.18)**
D(1994)	2.429	2.849	2.789	1.851
	(17.54)**	(25.99)**	(13.11)**	(9.47)**
D(1995)	3.116	2.930	3.419	2.578
	(22.18)**	(26.94)**	(15.78)**	(13.16)**
D(1996)	3.021	3.048	3.338	2.466
	(22.10)**	(28.06)**	(15.69)**	(12.96)**
D(1997)	3.344	3.224	3.629	2.793
	(23.92)**	(29.47)**	(16.44)**	(14.55)**
D(1998)	3.181	2.962	3.483	2.609
	(22.47)**	(27.05)**	(15.84)**	(13.37)**
D(1999)	2.393	2.536	2.931	1.683
	(17.96)**	(23.49)**	(13.96)**	(9.12)**
D(2000)	2.341	2.948	2.933	1.629
	(16.87)**	(26.80)**	(13.00)**	(8.57)**
D(Commerce)	0.149	0.102	0.176	0.146
	(3.57)**	(7.24)**	(2.82)**	(2.60)**
D(Service)	-0.071	-0.017	0.039	-0.136
	(1.48)	(0.99)	(0.50)	(2.18)*
Growth of quarterly GDP	-2.281	0.159	-2.289	-2.043
	(2.60)**	(0.51)	(1.83)	(1.65)
Constant	-5.597	-7.039	-2.212	-5.740
	(4.22)**	(17.71)**	(0.84)	(2.61)**
Observations	50755	69658	24917	25838
Pseudo-R^2	0.18	0.29	0.17	0.19

Table 4.4: Probit estimates for the approval of the first loan application. (a) based on applications with balance info, (b) based on all applications, the third column is based on observations with balance info and assets $\leq$ \$US 1,000, the fourth column is based on observations with balance info and assets > \$US 1,000. Sectors are relative to the production sector, branches relative to the main branch in La Paz.

Variable	Rejection with matching balance obs., N=1,490 Mean	Std.Dev.	All rejected applications, N=20,143 Mean	Std.Dev.	Approval, N=49,281 Mean	Std.Dev.
ln(amount appl. for)	6.106	*(1.192)*	6.073	*(1.190)*	5.650	*(0.956)*
liabilities/assets	0.059	*(0.175)*			0.030	*(0.123)*
ln(business age+1)	1.718	*(0.881)*			1.549	*(0.886)*
D(on black list)	0.027	*(0.163)*	0.018	*(0.134)*	4.10e-03	*(0.064)*
D(single)	0.239	*(0.427)*	0.719	*(0.449)*	0.197	*(0.398)*
D(female)	0.562	*(0.496)*			0.612	*(0.487)*
D(1992)	0.045	*(0.208)*	0.032	*(0.176)*	3.64e-04	*(0.019)*
D(1993)	0.074	*(0.262)*	0.009	*(0.094)*	0.026	*(0.159)*
D(1994)	0.107	*(0.309)*	0.040	*(0.196)*	0.101	*(0.302)*
D(1995)	0.091	*(0.287)*	0.118	*(0.323)*	0.165	*(0.371)*
D(1996)	0.147	*(0.354)*	0.154	*(0.361)*	0.219	*(0.414)*
D(1997)	0.088	*(0.284)*	0.104	*(0.306)*	0.166	*(0.372)*
D(1998)	0.105	*(0.306)*	0.145	*(0.352)*	0.149	*(0.356)*
D(1999)	0.225	*(0.418)*	0.279	*(0.449)*	0.118	*(0.323)*
D(2000)	0.118	*(0.323)*	0.119	*(0.324)*	0.055	*(0.228)*
D(Commerce)	0.474	*(0.499)*	0.503	*(0.500)*	0.540	*(0.498)*
D(Service)	0.260	*(0.439)*	0.247	*(0.431)*	0.175	*(0.380)*
D(Production)	0.266	*(0.442)*	0.250	*(0.433)*	0.284	*(0.451)*
D(Cochabamba)	0.110	*(0.314)*	0.179	*(0.383)*	0.105	*(0.306)*
D(La Paz)	0.467	*(0.499)*	0.407	*(0.491)*	0.664	*(0.472)*
D(Santa Cruz)	0.059	*(0.237)*	0.040	*(0.196)*	0.012	*(0.110)*
D(Sucre)	0.150	*(0.358)*	0.165	*(0.372)*	0.105	*(0.306)*
D(Trinidad)	0.101	*(0.302)*	0.086	*(0.280)*	0.042	*(0.200)*
D(Tarija)	0.112	*(0.315)*	0.122	*(0.328)*	0.072	*(0.259)*
Growth, quart.GDP	0.033	*(0.026)*	0.032	*(0.025)*	0.041	*(0.025)*

Table 4.5: Summary of means by approval of the first loan application.

	P(2nd balance obs.)		
	(all)	(≤1,000)	(>1,000)
ln(approved loan size)	0.350	0.612	0.091
	(4.50)**	(2.43)*	(0.61)
ln(approved loan size)^2	-0.023	-0.049	-0.002
	(3.14)**	(1.77)	(0.18)
ln(income)	0.025	0.014	0.025
	(1.02)	(0.41)	(0.74)
applied amount/approved amount	-0.017	-0.008	-0.029
	(2.94)**	(1.01)	(3.32)**
Latest payment:			
D(1 day)	0.025	-0.015	0.063
	(0.93)	(0.41)	(1.56)
D(2 days)	-0.199	-0.201	-0.194
	(5.14)**	(3.63)**	(3.58)**
D(3-5 days)	-0.236	-0.258	-0.212
	(8.71)**	(6.63)**	(5.55)**
D(6-10 days)	-0.791	-0.915	-0.667
	(29.13)**	(24.28)**	(16.88)**
D(11-29 days)	-1.602	-1.674	-1.525
	(58.05)**	(42.54)**	(39.19)**
D(≥ 30 days)	-2.613	-2.741	-2.516
	(50.86)**	(32.56)**	(38.35)**
length of loan in days	0.019	0.026	0.018
	(27.57)**	(17.13)**	(18.87)**
length in days^2	-8.62e-05	-1.48e-04	-7.95e-05
	(21.18)**	(10.74)**	(15.36)**
length in days^3	1.36e-07	3.38e-07	1.22e-07
	(15.49)**	(7.28)**	(11.63)**
length in days^4	-7.08e-11	-2.62e-10	-6.21e-11
	(12.04)**	(5.32)**	(9.20)**
D(female)	0.134	0.152	0.120
	(7.45)**	(5.83)**	(4.77)**
D(single)	-0.072	-0.114	-0.029
	(3.22)**	(3.69)**	(0.88)
ln(business age+1)	0.041	0.046	0.035
	(4.28)**	(3.48)**	(2.44)*
D(Commerce)	0.076	0.105	0.052
	(3.84)**	(3.82)**	(1.78)
D(Service)	0.024	0.046	0.005
	(0.92)	(1.21)	(0.12)
Growth of quarterly GDP	-0.697	-0.220	-1.249
	(1.88)	(0.43)	(2.31)*
Constant	-2.461	-3.071	-1.666
	(5.19)**	(3.42)**	(2.36)*
Observations	33323	17066	16257
Pseudo-R^2	0.23	0.22	0.23

Table 4.6: Probit estimates for the existence of a second balance information. The first column comprises all clients, the second those with assets below $US 1,000, the third those with assets above $US 1,000. Sample: First loan application in or before 1997 with a corresponding balance observation. Sectors are relative to the production sector. Dummy variables for years and branches have been included but are not reported here.

	No 2nd balance obs. N=8,647		With 2nd balance obs. N=24,684	
Variable	Mean	Std. Dev.	Mean	Std. Dev
ln(approved amount)	5.197	*(0.955)*	5.221	*(0.926)*
ln(income)	5.408	*(0.570)*	5.442	*(0.560)*
applied/aproved amount	1.820	*(1.437)*	1.730	*(1.390)*
maximum arrears	33.50	*(131.1)*	2.391	*(10.54)*
loan length in days	149.5	*(104.4)*	140.0	*(82.51)*
D(female)	0.585	*(0.493)*	0.636	*(0.481)*
D(single)	0.201	*(0.401)*	0.162	*(0.368)*
ln(business age+1)	1.562	*(0.937)*	1.620	*(0.918)*
D(1992)	0.002	*(0.040)*	0.000	*(0.013)*
D(1993)	0.047	*(0.211)*	0.035	*(0.184)*
D(1994)	0.130	*(0.336)*	0.156	*(0.363)*
D(1995)	0.209	*(0.406)*	0.255	*(0.436)*
D(1996)	0.343	*(0.475)*	0.317	*(0.465)*
D(1997)	0.270	*(0.444)*	0.237	*(0.425)*
D(Commerce)	0.506	*(0.500)*	0.572	*(0.495)*
D(Production)	0.186	*(0.390)*	0.149	*(0.357)*
D(Service)	0.308	*(0.462)*	0.279	*(0.448)*
D(Cochabamba)	0.144	*(0.351)*	0.107	*(0.309)*
D(La Paz)	0.657	*(0.475)*	0.693	*(0.461)*
D(Sucre)	0.099	*(0.298)*	0.111	*(0.314)*
D(Trinidad)	0.014	*(0.118)*	0.016	*(0.125)*
D(Tarija)	0.086	*(0.280)*	0.074	*(0.261)*
Growth of quart.GDP	0.046	*(0.023)*	0.045	*(0.023)*

Table 4.7: Summary statistics for approved loan applications by existence of a second balance observation, given that the first loan application was approved no later than 1997.

	Dependent variable: growth in assets					
	Commerce		Production		Service	
	(a)	(b)	(a)	(b)	(a)	(b)
ln(assets)	-2.332	-2.346	-3.152	-3.152	-1.156	-1.161
	(3.38)**	(3.40)**	(3.51)**	(3.52)**	(0.97)	(0.97)
ln(assets)^2	0.138	0.139	0.293	0.294	0.066	0.066
	(2.23)*	(2.25)*	(3.17)**	(3.17)**	(0.49)	(0.49)
ln(assets)^3	-0.005	-0.005	-0.011	-0.011	-0.003	-0.003
	(2.03)*	(2.01)*	(2.83)**	(2.81)**	(0.51)	(0.51)
ln(income)	0.103	0.103	0.078	0.086	0.240	0.241
	(4.84)**	(4.71)**	(2.66)**	(2.92)**	(4.21)**	(4.18)**
NUMPRIAP	-0.034	-0.035	-0.244	-0.256	0.151	0.151
	(0.49)	(0.50)	(2.65)**	(2.77)**	(0.90)	(0.90)
NUMPRIAP*ln(assets)	0.018	0.018	0.041	0.043	-0.012	-0.012
	(1.99)*	(2.00)*	(3.29)**	(3.43)**	(0.54)	(0.54)
APRISIZE / assets	1.314	1.301	0.714	0.717	1.240	1.244
	(4.08)**	(4.06)**	(3.48)**	(3.50)**	(4.80)**	(4.80)**
ln(APRISIZE)	0.165	0.172	0.162	0.163	0.043	0.041
	(2.73)**	(2.89)**	(3.96)**	(3.98)**	(0.83)	(0.79)
ln(time)	-0.392	-0.388	-0.267	-0.255	-0.274	-0.274
	(18.72)**	(19.12)**	(11.55)**	(11.90)**	(7.38)**	(7.86)**
D(female)	-0.042	-0.042	-0.029	-0.029	-0.036	-0.036
	(2.58)*	(2.54)*	(1.63)	(1.62)	(1.18)	(1.16)
ln(age)	-0.505	-0.503	-0.702	-0.696	-0.405	-0.406
	(3.21)**	(3.20)**	(2.77)**	(2.74)**	(1.25)	(1.25)
ln(age)*ln(assets)	0.055	0.055	0.078	0.079	0.033	0.034
	(2.42)*	(2.41)*	(2.19)*	(2.22)*	(0.76)	(0.77)
D(Cochabamba)	0.034	0.047	0.017	0.029	-0.002	-0.007
	(1.24)	(1.88)	(0.51)	(0.92)	(0.03)	(0.11)
D(Sucre)	0.091	0.108	0.041	0.060	0.190	0.186
	(2.79)**	(3.65)**	(0.82)	(1.26)	(3.57)**	(4.05)**
D(Trinidad)	0.022	0.047	0.099	0.132	0.018	0.013
	(0.44)	(1.11)	(0.91)	(1.26)	(0.23)	(0.20)
D(Tarija)	-0.063	-0.042	-0.086	-0.061	-0.050	-0.055
	(2.24)*	(2.26)*	(2.15)*	(1.78)	(0.91)	(1.49)
λ_1	0.069	-0.007	0.151	0.081	-0.007	0.013
	(0.84)	(0.22)	(2.01)*	(1.95)	(0.05)	(0.23)
λ_2	-0.114	-0.110	-0.150	-0.147	-0.098	-0.099
	(3.44)**	(3.31)**	(4.35)**	(4.28)**	(1.86)	(1.90)
Constant	13.051	13.088	11.662	11.558	7.365	7.381
	(3.25)**	(3.26)**	(2.73)**	(2.72)**	(1.31)	(1.31)
Observations	14072	14072	6869	6869	3674	3674
R-squared	0.16	0.16	0.17	0.17	0.17	0.17

Table 4.8: Estimated growth in assets by sectors. Robust t-statistics are in parentheses, * indicates significant at 5%; ** significant at 1%. D(.) indicates a dummy variable, dummies for quarters, branches, years, GNP growth and interactions of year dummies and log assets have been included but are not reported here. Branches are relative to the main branch in La Paz. Columns (a) correspond to approval estimates based on balance information (column (a) in table 4.4), columns (b) correspond to approval estimates based on data available for all applicants (column (b) in table 4.4).

	Dep. variable: growth in fixed assets	
	(a)	(b)
ln(fixed assets)	-0.857	-0.839
	(5.30)**	(5.19)**
ln(fixed assets)^2	0.110	0.104
	(3.48)**	(3.29)**
ln(fixed assets)^3	-0.006	-0.006
	(3.16)**	(2.92)**
ln(income)	0.102	0.120
	(1.48)	(1.68)
NUMPRIAP	-0.329	-0.346
	(1.87)	(1.96)
NUMPRIAP*ln(assets)	0.043	0.046
	(1.85)	(1.95)
APRISIZE / assets	0.607	0.612
	(1.80)	(1.80)
ln(APRISIZE)	0.176	0.191
	(2.68)**	(2.88)**
ln(time)	-0.113	-0.078
	(2.02)*	(1.51)
D(female)	-0.122	-0.122
	(2.45)*	(2.45)*
ln(age)	-0.239	-0.226
	(2.63)**	(2.44)*
D(Cochabamba)	0.083	0.152
	(0.85)	(1.73)
D(Sucre)	0.198	0.288
	(1.71)	(2.58)**
D(Trinidad)	0.398	0.549
	(2.00)*	(3.05)**
D(Tarija)	0.030	0.145
	(0.31)	(1.81)
Growth of quart.GDP	-3.975	-3.886
	(3.33)**	(3.26)**
λ_1	0.459	0.086
	(2.16)*	(0.74)
λ_2	-0.241	-0.221
	(2.98)**	(2.76)**
Constant	1.409	1.444
	(0.62)	(0.63)
Observations	6873	6873
R-squared	0.09	0.09

Table 4.9: Estimated growth in fixed assets for the production sector. * indicates significant at 5%; ** significant at 1%. D(.) indicates a dummy variable, dummies for quarters, years, and interaction terms of years and log assets have been included but are not reported here. Branches are relative to the main branch in La Paz. Column (a) corresponds to approval estimates based on balance information (column (a) in table 4.4), column (b) corresponds to approval estimates based on data available for all applicants (column (b) in table 4.4).

	Dependent variable: ln(sales), Commerce Sector					
	(a)	(b)	(OLS)	(a)	(b)	(OLS)
ln(assets)	0.577	0.410	0.407	0.598	0.420	0.417
	(11.00)**	(7.88)**	(7.78)**	(11.40)**	(8.07)**	(7.97)**
$\ln(\text{assets})^2/2$	0.031	0.061	0.064	0.027	0.059	0.062
	(6.55)**	(13.62)**	(14.32)**	(5.62)**	(13.04)**	(13.76)**
ln(Num.employees)	0.586	0.622	0.616	0.586	0.623	0.618
	(4.22)**	(4.34)**	(4.30)**	(4.20)**	(4.32)**	(4.29)**
$\ln(\text{employees})^2/2$	0.002	-0.019	-0.014	0.002	-0.019	-0.014
	(0.09)	(0.68)	(0.49)	(0.07)	(0.68)	(0.51)
ln(assets)*ln(empl.)	-0.079	-0.079	-0.078	-0.079	-0.079	-0.079
	(4.83)**	(4.59)**	(4.59)**	(4.80)**	(4.56)**	(4.57)**
D	0.037	0.046	0.014	0.057	-0.056	-0.055
	(3.74)**	(4.55)**	(2.09)*	(0.85)	(0.84)	(0.82)
$D \cdot$ ln(assets)				0.037	0.025	0.026
				(5.87)**	(4.00)**	(4.11)**
$D \cdot$ ln(time)				-0.056	-0.016	-0.022
				(5.10)**	(1.46)	(2.04)*
liabilities/assets	0.416	0.466	0.463	0.409	0.464	0.460
	(8.83)**	(10.03)**	(9.95)**	(8.72)**	(9.99)**	(9.91)**
D(female)	-0.083	-0.038	-0.029	-0.078	-0.033	-0.024
	(1.71)	(0.77)	(0.59)	(1.61)	(0.66)	(0.49)
D(female)*ln(assets)	0.003	-0.004	-0.004	0.002	-0.004	-0.005
	(0.40)	(0.53)	(0.65)	(0.32)	(0.62)	(0.74)
ln(age)	0.331	0.374	0.377	0.335	0.378	0.381
	(4.46)**	(5.02)**	(5.05)**	(4.53)**	(5.08)**	(5.11)**
ln(age)*ln(assets)	-0.060	-0.063	-0.067	-0.060	-0.063	-0.067
	(5.50)**	(5.75)**	(6.06)**	(5.54)**	(5.79)**	(6.10)**
ln(busage)	0.181	0.167	0.167	0.195	0.176	0.176
	(6.80)**	(6.27)**	(6.23)**	(7.31)**	(6.58)**	(6.56)**
ln(busage)*ln(assets)	-0.013	-0.009	-0.009	-0.014	-0.010	-0.010
	(3.25)**	(2.35)*	(2.31)*	(3.70)**	(2.63)**	(2.61)**
D(Cochabamba)	-0.083	0.045	0.068	-0.095	0.042	0.066
	(6.54)**	(3.98)**	(6.28)**	(7.38)**	(3.73)**	(6.05)**
D(Sucre)	-0.525	-0.344	-0.320	-0.532	-0.345	-0.321
	(38.64)**	(32.18)**	(30.76)**	(38.95)**	(32.22)**	(30.79)**
D(Trinidad)	-0.545	-0.249	-0.221	-0.559	-0.250	-0.222
	(21.26)**	(11.73)**	(10.48)**	(21.70)**	(11.78)**	(10.54)**
D(Tarija)	-0.308	-0.075	-0.046	-0.319	-0.076	-0.047
	(20.55)**	(7.05)**	(4.45)**	(21.04)**	(7.16)**	(4.53)**
λ_1	0.920	0.124		0.952	0.125	
	(23.78)**	(10.65)**		(24.23)**	(10.75)**	
$D \cdot \lambda_2$	-0.131	-0.093		-0.097	-0.077	
	(5.36)**	(3.81)**		(3.93)**	(3.09)**	
Constant	1.871	2.777	2.834	1.789	2.749	2.807
	(6.30)**	(9.41)**	(9.57)**	(6.02)**	(9.32)**	(9.48)**
Observations	35796	36805	36805	35796	36805	36805
R-squared	0.63	0.62	0.62	0.63	0.63	0.62

Table 4.10: Estimated translog production function for the commerce sector. Robust t-statistics are in parentheses. * indicates significant at 5%; ** significant at 1%. D(.) indicates a dummy variable, dummies for years and quarters and GNP growth variables have been included but are not reported here. Branches are relative to the main branch in La Paz. Columns (a) correspond to approval estimates based on balance information (column (a) in table 4.4), columns (b) correspond to approval estimates based on data available for all applicants (column (b) in table 4.4). Columns (OLS) report OLS estimates based on the same sample as columns (b).

	Dependent variable: ln(sales), Production Sector					
	(a)	(b)	(OLS)	(a)	(b)	(OLS)
ln(assets)	0.761	0.662	0.662	0.774	0.672	0.672
	(8.78)**	(7.81)**	(7.80)**	(8.92)**	(7.92)**	(7.93)**
ln(assets)^2/2	0.009	0.029	0.029	0.007	0.027	0.027
	(1.17)	(4.12)**	(4.15)**	(0.85)	(3.83)**	(3.84)**
ln(Num.employees)	0.339	0.296	0.290	0.339	0.295	0.289
	(4.33)**	(3.76)**	(3.68)**	(4.33)**	(3.74)**	(3.67)**
ln(employees)^2/2	-0.027	-0.031	-0.029	-0.027	-0.030	-0.028
	(0.78)	(0.88)	(0.82)	(0.76)	(0.83)	(0.78)
ln(assets)*ln(empl.)	-0.026	-0.018	-0.017	-0.026	-0.018	-0.017
	(2.47)*	(1.65)	(1.63)	(2.47)*	(1.65)	(1.63)
D	0.119	0.127	0.059	0.040	-0.105	-0.125
	(8.27)**	(8.80)**	(6.46)**	(0.45)	(1.17)	(1.40)
D · ln(assets)				0.027	0.023	0.028
				(2.95)**	(2.51)*	(3.03)**
D · ln(time)				-0.024	0.012	-0.003
				(1.63)	(0.87)	(0.22)
liabilities/assets	0.030	0.162	0.158	0.026	0.161	0.157
	(0.29)	(1.96)*	(1.97)*	(0.25)	(1.95)	(1.96)*
D(female)	-0.278	-0.261	-0.262	-0.274	-0.257	-0.257
	(4.41)**	(4.14)**	(4.15)**	(4.34)**	(4.07)**	(4.07)**
D(female)*ln(assets)	0.031	0.029	0.030	0.031	0.029	0.029
	(3.51)**	(3.29)**	(3.34)**	(3.45)**	(3.23)**	(3.26)**
ln(age)	0.445	0.501	0.488	0.450	0.503	0.492
	(3.45)**	(3.89)**	(3.79)**	(3.49)**	(3.91)**	(3.82)**
ln(age)*ln(assets)	-0.080	-0.087	-0.087	-0.081	-0.087	-0.087
	(4.32)**	(4.70)**	(4.66)**	(4.36)**	(4.71)**	(4.68)**
ln(business age)	0.058	0.066	0.062	0.068	0.073	0.072
	(1.48)	(1.68)	(1.60)	(1.75)	(1.87)	(1.84)
ln(busage)*ln(assets)	-0.005	-0.005	-0.004	-0.006	-0.006	-0.006
	(0.83)	(0.84)	(0.75)	(1.06)	(1.04)	(0.99)
D(Cochabamba)	-0.074	0.035	0.039	-0.079	0.035	0.038
	(2.93)**	(2.02)*	(2.29)*	(3.06)**	(2.03)*	(2.25)*
D(Sucre)	-0.479	-0.340	-0.331	-0.482	-0.341	-0.331
	(15.70)**	(16.57)**	(16.24)**	(15.61)**	(16.60)**	(16.24)**
D(Trinidad)	-0.160	0.049	0.061	-0.169	0.047	0.058
	(2.72)**	(1.07)	(1.34)	(2.84)**	(1.03)	(1.27)
D(Tarija)	-0.079	0.100	0.110	-0.083	0.099	0.110
	(2.08)*	(4.22)**	(4.77)**	(2.16)*	(4.18)**	(4.75)**
λ_1	0.597	0.045		0.610	0.046	
	(6.56)**	(2.61)**		(6.55)**	(2.67)**	
$D \cdot \lambda_2$	-0.220	-0.189		-0.207	-0.187	
	(6.85)**	(5.87)**		(6.38)**	(5.76)**	
Constant	1.173	1.740	1.783	1.125	1.725	1.762
	(2.31)*	(3.52)**	(3.60)**	(2.21)*	(3.49)**	(3.56)**
Observations	18233	18644	18644	18233	18644	18644
R-squared	0.55	0.55	0.55	0.56	0.55	0.55

Table 4.11: Estimated translog production function for the production sector. Robust t-statistics are in parentheses, * indicates significant at 5%; ** significant at 1%. D(.) indicates a dummy variable, dummies for quarters and GNP growth have been included but are not reported here. Branches are relative to the main branch in La Paz. Columns (a) correspond to approval estimates based on balance information (column (a) in table 4.4), columns (b) correspond to approval estimates based on data available for all applicants (column (b) in table 4.4), and columns (OLS) are OLS estimates.

	Dependent variable: ln(sales), Service Sector					
	(a)	(b)	(OLS)	(a)	(b)	(OLS)
ln(assets)	0.407	0.271	0.290	0.427	0.277	0.297
	(4.16)**	(2.80)**	(3.00)**	(4.34)**	(2.85)**	(3.06)**
$\ln(\text{assets})^2/2$	-0.002	0.024	0.025	-0.006	0.023	0.024
	(0.26)	(2.67)**	(2.74)**	(0.63)	(2.54)*	(2.59)**
ln(Num.employees)	0.176	0.128	0.107	0.185	0.131	0.112
	(0.92)	(0.66)	(0.55)	(0.97)	(0.67)	(0.57)
$\ln(\text{employees})^2/2$	-0.027	-0.046	-0.046	-0.030	-0.047	-0.048
	(0.79)	(1.24)	(1.23)	(0.86)	(1.27)	(1.28)
ln(assets)*ln(empl.)	0.007	0.019	0.021	0.006	0.019	0.021
	(0.29)	(0.80)	(0.91)	(0.26)	(0.79)	(0.90)
D	0.152	0.161	0.089	0.371	0.264	0.263
	(6.82)**	(7.23)**	(6.15)**	(2.63)**	(1.90)	(1.89)
$D \cdot$ ln(assets)				0.034	0.012	0.016
				(2.73)**	(0.94)	(1.27)
$D \cdot$ ln(time)				-0.092	-0.037	-0.055
				(3.67)**	(1.51)	(2.31)*
liabilities/assets	0.783	0.890	0.883	0.772	0.887	0.879
	(7.40)**	(8.37)**	(8.29)**	(7.30)**	(8.34)**	(8.25)**
D(female)	0.557	0.621	0.625	0.569	0.626	0.634
	(6.47)**	(7.20)**	(7.25)**	(6.59)**	(7.24)**	(7.33)**
D(female)*ln(assets)	-0.045	-0.052	-0.053	-0.046	-0.053	-0.054
	(3.73)**	(4.36)**	(4.40)**	(3.83)**	(4.41)**	(4.48)**
ln(age)	0.423	0.467	0.485	0.427	0.470	0.488
	(2.85)**	(3.16)**	(3.28)**	(2.88)**	(3.18)**	(3.30)**
ln(age)*ln(assets)	-0.037	-0.041	-0.047	-0.038	-0.041	-0.047
	(1.77)	(1.94)	(2.22)*	(1.79)	(1.95)	(2.23)*
ln(business age)	-0.132	-0.126	-0.131	-0.117	-0.121	-0.124
	(2.54)*	(2.44)*	(2.54)*	(2.24)*	(2.32)*	(2.38)*
ln(busage)*ln(assets)	0.025	0.027	0.028	0.023	0.027	0.027
	(3.39)**	(3.73)**	(3.84)**	(3.12)**	(3.63)**	(3.71)**
D(Cochabamba)	-0.118	0.097	0.115	-0.134	0.094	0.111
	(3.76)**	(3.49)**	(4.24)**	(4.24)**	(3.37)**	(4.08)**
D(Sucre)	-0.723	-0.455	-0.436	-0.732	-0.454	-0.435
	(26.92)**	(21.78)**	(21.12)**	(27.15)**	(21.76)**	(21.10)**
D(Trinidad)	-0.623	-0.169	-0.144	-0.645	-0.171	-0.148
	(13.63)**	(4.95)**	(4.23)**	(14.00)**	(5.01)**	(4.35)**
D(Tarija)	-0.432	-0.076	-0.052	-0.444	-0.076	-0.051
	(13.39)**	(3.36)**	(2.37)*	(13.67)**	(3.34)**	(2.35)*
λ_1	1.148	0.098		1.184	0.098	
	(16.09)**	(3.96)**		(16.47)**	(3.95)**	
$D \cdot \lambda_2$	-0.207	-0.191		-0.168	-0.175	
	(4.47)**	(4.17)**		(3.56)**	(3.73)**	
Constant	2.075	3.108	3.074	1.975	3.083	3.041
	(3.60)**	(5.48)**	(5.42)**	(3.42)**	(5.43)**	(5.36)**
Observations	10191	10574	10574	10191	10574	10574
R-squared	0.32	0.30	0.30	0.32	0.30	0.30

Table 4.12: Estimated translog production function for the service sector. Robust t-statistics are in parentheses, * indicates significant at 5%; ** significant at 1%. D(.) indicates a dummy variable, dummies for quarters and GNP growth have been included but are not reported here. Branches are relative to the main branch in La Paz. Columns (a) correspond to approval estimates based on balance information (column (a) in table 4.4), columns (b) correspond to approval estimates based on data available for all applicants (column (b) in table 4.4), and columns (OLS) are OLS estimates.

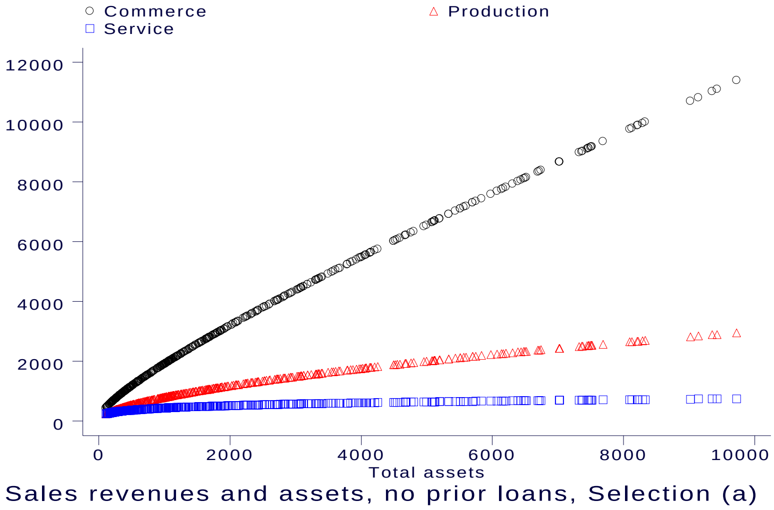

Figure 4.8: Predicted ratio of monthly sales revenues over assets by business sector for a male client in La Paz in 1996 with 30 years of age and a 4 year old business.

Chapter 5

Microfinance in Times of Crisis: The Effects of Competition, Rising Indebtedness, and Economic Crisis on Repayment Behavior

5.1 Introduction

Most developing countries have a large informal sector, constituted of small unregistered businesses. The majority of these micro-enterprises suffers from an inadequately low level of capital since their owners do not have access to the formal banking sector.[1] Since the late 1970's, development policy has increasingly taken recourse to microfinance to improve the access to financial services for poor households. Compared to previous attempts to provide credit to the poor, the novelty of microfinance consists firstly in the use of new incentive mechanisms such as group loans or the choice of collateral based on the borrower's subjective valuation, and secondly in the attempt to cover costs through high interest rates.[2]

In recent years, an increasing number of microfinance institutions finds it hard to maintain high repayment rates. While many microfinance institutions were the

[1]The term micro-enterprise refers to small businesses in the informal sector. They encompass a wide range of enterprises from small family run trade shops to production businesses with many employees and high revenues. Micro-enterprises are typically characterized in terms of number of employees or assets instead of legal status. The Inter-American Development Bank (IADB), for example, defines a micro-enterprise as having no more than 10 workers and total assets below $US 20.000 (Orlando and Pollack 2000).

[2]For more information about microfinance see Morduch (2000).

sole source of reasonably priced loans in their early years of operation, today clients frequently can choose between institutions. Since funds are not as scarce as they used to be, the incentives to repay on time and thus to remain in good standing with the institutions have decreased. In a few areas, notably in Bangladesh and Bolivia, the microfinance market is close to saturation. Matin (2001) and Chaudhury and Matin (2001) report that an increasing number of households takes loans from multiple institutions in Bangladesh and that the repayment performance declines. The authors report estimates of market coverage between 43% and 59%. For Bolivia, Rhyne (2001, pp. 19, 31) estimates that between a quarter and a third of all micro-enterprises obtain microfinance loans. Besides the high supply of loans in the Bolivian microfinance market, the economic environment has been characterized by severe difficulties since 1998. Consumer credit companies (most of which are out of the market today) have distributed loans to many micro-entrepreneurs. These borrowers had increasingly high debt levels and repayment obligations, which they frequently could not fulfill. Since the end of 1998, the economy has slowed down with negative growth in 1999 and a low level of economic activity since. These developments together led to a crisis in microfinance lending which culminated on July 2nd, 2001, when a group of people from a debtor association took employees from the superintendency of banks as hostages and demanded debt forgiveness.[3]

The Bolivian microfinance institutions have faced a strong increase in late payments during these years. Between 1996 and 2000 the percentage of overdue capital rose from 2.6% to 12.3% for BancoSol and from 3.97% to 7.7% for Caja Los Andes, to name two of the largest Bolivian microlenders (ASOFIN 2000). After its strong initial success, microfinance seems to have reached a level where the institutions need to develop new strategies to maintain their good performance in a more competitive environment. If microfinance is to offer long-term services, it has to prove that it works in non-monopolistic environments and that it can maintain high repayment incentives even in the face of increasing saturation and competition.

In the face of these developments, it is essential for microfinance institutions to

[3]The recent developments in Bolivia are summarized in Von Stauffenberg (2001) and a detailed report has been provided by the BBC on July 4th, 2001. A summary of BBC's reports is listed at *http://nt1.ids.ac.uk/cgap/html/bolbbchl.htm*.

select their clients cautiously and to provide sufficient incentives for repayment. The analysis in the next sections analyzes the repayment behavior of clients from one bank in Bolivia, Caja Los Andes. Our focus lies on the discussion of the increase in late payments in recent years. Can we attribute this increase to the effects of the economic crisis beginning in late 1998, the over-indebtedness of many clients, or the rising competition? Which of these factors dominates? From a statistical point of view, the changes in the microfinance environment discussed above lend themselves readily to an econometric analysis. Most of the important changes are exogenous and their effects can be identified since there is variation over time and also between different geographic locations in Bolivia.

The paper continues with a discussion of repayment incentives and related literature in section 5.2. Section 5.3 presents the theoretical model underlying our analysis and section 5.4 briefly describes the data set used. Econometric issues are discussed in section 5.5. Section 5.6 presents the results and section 5.7 discusses their implications.

5.2 Overview of Repayment Incentives and Related Literature

Why should clients repay their loans on time? When analyzing repayment determinants for microfinance loans, the first step lies in the analysis of repayment incentives which largely depend on the terms and conditions of the loan. For individual loans, we can single out four major incentives for a timely repayment. Firstly, clients lose their guarantees. All loans are secured by chattel items and/or personal guarantees. Larger loans also can be secured through mortgages. If the client does not repay the loan, the chattel items are confiscated. The chattel items are selected based on the client's valuation and not on the resale value of these items which increases the client's repayment incentives.

Secondly, if the client does not repay the loan, he loses access to future loans. Caja Los Andes does not grant consecutive loans for defaulting clients and for clients with frequent late payments. In addition, the client obtains a bad credit record with the Bolivian banking supervisory authority. All registered banks and an increasing number of non-regulated microlenders can access this information and will not grant loans to this client in the future. As a consequence, the client will have to use informal sources or moneylenders for future loans. These loans tend to be more expensive: moneylenders charge higher interest rates and informal loans tend to have additional social costs.

Thirdly, conditions of the loan improve for clients with timely repayment. Loan sizes increase and repayment schedules become more flexible. Eventually, clients can obtain an automatic credit line. These improved conditions reduce the non-pecuniary costs of the loan for the client. Finally, the client's income needs to be sufficiently high to enable him to repay the loan on time. If the installments are too high or if his revenues are lower than expected he cannot repay the loan on time. If he obtains lower than expected revenues from his business, alternative sources of income are crucial for his ability to repay.

There is a body of literature that asks whether it can be optimal for a bank to behave as described above. In particular, when can it be optimal to offer relatively small loans to new clients although larger loans tend to be more profitable? When can it be optimal for a bank not to grant consecutive loans if the first loan is not repaid fully or not on time?

A characteristic crucial for the analysis of microfinance markets is the difficulty of finding adequate collateral. Ghosh and Ray (1999) analyze a market for loans where there is no collateral and there are no credit histories. That is, the bank has no information about new clients. They show that it is optimal for banks to distinguish between old and new clients and thus generate "inside" reputation mechanisms. New clients are offered small loans to test their repayment behavior. Once the first loan has been repaid, clients are offered larger loans at better conditions making it desirable to repay each loan in order not to lose the preferred client status. This behavior closely corresponds to patterns observed in microfinance markets.

Due to the relatively widespread availability of credit records in the Bolivian microfinance market, reputation effects play a considerable role beyond the borrower-lender relationship. In an early analysis of the provision of loans without collateral, Allen (1981) shows that reputation effects lead to the existence of incentive compatible lending contracts even in the absence of collateral. The model assumes that the termination threat (i. e. the threat not to extend another loan if the current loan is not repaid) is credible due to reputation building by the lender, for example. As a consequence, a borrower repays as long as the present value of future loans is higher than the current payment due.

Reputation building by the lender seems to play an important role in microfinance markets. In a study of five microlenders that distribute individual loans only, Churchill (1999) finds that the signals given to other borrowers are among the most important reasons for the banks not to extend a new loan to defaulting or late paying clients. The credibility assumption, however, is not as innocuous as it might seem from these examples since the bank might be able to increase its profits from future loans and, a priori, it is unclear which effect dominates. Lenders would like to pre-commit themselves not to extend the credit limit if the borrower cannot repay, since the expectation of a loan extension changes the borrower's behavior even in the case of full information as shown in Hellwig (1977). However, it is generally not optimal for them to follow this policy once the client has defaulted. At this point, the bank might find it optimal to distribute a new loan to this client to recover parts of the old loan, leading to a commitment problem. In a related analysis, Gromb (2001) shows that the termination threat is credible only if the lender makes zero profit from renegotiated loans. Hellwig mentions that in all likelihood financial markets develop institutions that allow the lender to pre-commit on credit limits. The Bolivian banking supervisory authority (Superintendencia de Bancos) serves such a need. The rescheduling of loans, for example, was allowed only if the lender set aside additional reserve funds.[4] This restriction made reschedulings more costly and made the bank's commitment not to reschedule more credible.

The special situation of microlenders bears similarities to international lending. Given that legal enforcement mechanisms are difficult to use and collateral is hard

[4]See Rhyne (2001, p. 150). This restriction was temporarily softened in 1999/2000 after the economic crisis led to severe repayment problems.

to seize, Eaton, Gersovitz, and Stiglitz (1986) argue that loan contracts are offered only if there are other strong enforcement mechanisms, such as reputation effects on the side of the borrower. That is, if the borrower loses future access to these funds in the case of default. If this is not the case, the bank cannot generate sufficient incentives to ensure a repayment of the loan. The limitations of reputation effects are also highlighted in Bulow and Rogoff (1989), who show that reputation effects alone are not sufficient to generate repayment incentives when the borrower has access to sufficiently diversified investment opportunities.

The microfinance market has seen an enormous capital inflow during the last 15 years.[5] When one applies the models in Holmström and Tirole (1997) and Bolton and Freixas (2000) to this situation, they predict the provision of smaller and riskier loans. In the beginning, an intermediary can choose high-return, low-risk clients. If his funds increase, however, these clients eventually are served and he also considers clients with higher risks. Thus, the increasing supply of funds to the microfinance intermediaries should lead to increasing default rates because of a change in the client structure towards riskier clients. In addition, the increasing availability of microfinance loans leads to competition and affects the client's outside options. Villas-Boas and Schmidt-Mohr (1999), for example, show that increasing competition can lead to rising collateral requirements.

While these and other studies are mostly concerned with the existence and design of incentive compatible contracts, our approach differs in that we ask what determines whether or not a client repays, given the credit contract. Based on Bolton and Sharfstein (1990), Armendáriz de Aghion and Morduch (2000) analyze incentive effects for individual microfinance loans. They show the importance of collateral and social sanctions and provide an example for a mechanism through which a credible "non-refinancing threat" leads to a higher effort level of the client. In addition, regular repayment schedules as used in most microfinance institutions are shown to

[5]See Morduch (2000) for an overview. The portfolio of the microfinance institutions in La Paz, Bolivia, for example, has approximately increased from $US 10 in 1992 to $US 80 in 2000 per capita (working age population). Source: calculations based on information from ASOFIN (December 1999, table 2.19) and population data from the Instituto Nacional de Estadística (INE), La Paz, *http://www.ine.gov.bo*.

work as a disciplining device and, in addition, provide continuous information on the clients' ability and willingness to repay the loan. Eaton, Gersovitz, and Stiglitz (1986) argue that due to the absence of legal enforcement the bank has to generate very high incentive effects. These, in turn, reduce the importance of moral hazard and adverse selection (relative to standard banking).

While there are a number of theoretical studies on repayment behavior, empirical evidence is scarce due to a lack of adequate data. From a number of interviews with clients who have individual microfinance loans, Churchill (1999) finds that the continued access to future loans serves as the most important repayment incentive. Thus, banks generate high incentives for timely repayment when making future loans contingent on good repayment performance. Schreiner (1999) estimates the probability of high arrears for a Bolivian microlender (arrears = days overdue). Among other things, he finds a significant influence of business sectors, loan sizes (with larger loans being more likely to have high arrears), and in particular of past repayment behavior. Chaudhury and Matin (2001) describe a sample of Bangladesh households, finding that households with multiple loans at the same time tend to have lower repayment rates than others. Greene (1998) analyzes data from a credit card company and estimates default probabilities with a special emphasis on selection issues. He finds that when the estimates are based on clients only and selection issues are ignored, defaults are underestimated.

5.3 Theory

For a better understanding of the contractual obligations and incentive effects at work, this section begins with the description of a typical loan contract. If a loan application has been approved, the client and the bank agree on the number and frequency of the scheduled repayments. These payments typically are of a fixed amount and their size depends on the client's repayment capacity. Suppose, for example, a client obtains a loan of $US 1,000 with a monthly interest rate of 2.5%. A typical repayment schedule then would consist of ten monthly installments, that is, the client has to make ten payments of $114.26 and pays total interest of $142.59. The first payment is due 30 days after the client has taken out the loan.

If payments are late, clients have to pay a penalty in form of higher interest rates. If the client has not paid after a few days, his loan officer makes a visit, demands payment, and delivers an official letter reminding of the outstanding payment. In extraordinary circumstances, the loan officer might grant a postponement of payments or a rescheduling (this happens for 1.1% of all loans only). If the payments are overdue by more than 30 days and no postponement or rescheduling has taken place, the bank begins to collect collateral or takes the loan to court. In addition, the credit record is sent to the banking supervisory authority and the bank does not grant future loans.

To formalize the client's decision whether or not to repay a loan on time consider the following setting: A borrower has an outstanding loan of size L and non-business income of size V. He consumes an amount c and invests the remaining amount. The investment requires personal effort e and yields a return $g(V-L-c, e, A)$ if successful, where A represents idiosyncratic characteristics that determine productivity. The success probability is given by π and the client learns if he will have success before he chooses his effort and consumption levels.[6] We assume that $g(\cdot)$ is increasing in all arguments with decreasing returns in the first two arguments. Effort is costly in terms of utility as represented by $h(e)$ with $h_e > 0$ and $h_{ee} > 0$. For simplicity, we assume that the borrower has to repay all of the loan plus interest at the end of the period, amounting to $(1+r)L$.[7] If the loan is not paid on time, the client has to pay a penalty P. This penalty can consist of higher interest rates if the borrower pays late or of collateral seizure if he fails to pay at all.

The borrower obtains repeat loans from the same lender only if he has paid back the first loan on time. Otherwise, the lender does not grant further loans and the borrower obtains a bad credit record. From then on, he can obtain a loan from other lenders at higher interest rates only, if at all. Let B denote the future benefits from

[6]Extending the model to allow for an influence of effort on the success probability does not change the main results.

[7]The timing of scheduled payments plays an important role for the bank since late payments give an early signal of repayment problems. The incentive effects of these payments, however, are not fundamentally different from loans which have to be repaid in one lump-sum and we disregard the differences here. For an explicit analysis of frequently scheduled repayments, see Armendáriz de Aghion and Morduch (2000).

timely repayment. What determines the size of these benefits? First of all, the more widespread the availability of credit records, δ, the higher the barrier to obtain future loans. The higher the supply of loans and the competition among microlenders, the easier it is to obtain an alternative loan and the lower the benefits. Let ζ capture this effect. There may be size effects with poorer households being able to borrow from moneylenders instead. Let initial wealth W and non-business income V represent this effect. In addition, the firm's leverage, λ, and current loan size L determine to what extent the client needs consecutive loans to maintain the scale of his business. Finally, there may be idiosyncratic characteristics, A. These can be gender, business sector, or location which significantly determine alternative borrowing possibilities and thus the costs of losing access to future loans. In addition, A captures the client's subjective valuation of paying on time.

To simplify notation we denote everything in present value terms and abstract from discounting. In addition, let $v_{RS} = v(V+g(V+L-c,e,A)-(1+r)L)$ denote the second period utility when the client repays and is successful, $v_{RF} = v(V-(1+r)L)$ when he repays and has a business failure, $v_{DS} = v(V+g(V+L-c,e,A)-P)$ when he defaults and is successful, and $v_{DF} = v(V-P)$ when he defaults and has a business failure. The utilities from repayment (U^R) and default (U^D) are given by

$$\begin{aligned} U^R &= u(c) + \pi \cdot v_{RS} + (1-\pi) \cdot v_{RF} - h(e) + B(V,W,A,\delta,\zeta,\lambda,L) \quad \text{and} \\ U^D &= u(c) + \pi \cdot v_{DS} + (1-\pi) \cdot v_{DF} - h(e) \quad . \end{aligned} \tag{5.1}$$

From a direct comparison of U^R and U^D one can see that clients will choose to repay if interest rates are low, if the penalty P is high, if the future benefit from continued access to loans B is large, or if clients have relatively high returns on their investments. This is because marginal utility is decreasing and the difference between $(1+r)L$ and P in terms of utility is lower at higher values for $g(\cdot)$.

We can divide clients into three different groups depending on their optimal repayment behavior.[8]

$$\text{Optimal policy} = \begin{cases} \text{always repay} & B \;\geq\; \overline{B}(V,P,r,L) \\ \text{repay if successful} & \overline{B}(V,P,r,L) \;>\; B \;\geq\; \underline{B}(V,P,r,L,A) \\ \text{never repay} & \underline{B}(V,P,r,L,A) \;>\; B \quad , \end{cases} \tag{5.2}$$

where $\overline{B}(V,P,r,L) = v_{DF} - v_{RF}$ and $\underline{B}(V,P,r,L,A) = v_{DS} - v_{RS}$, where v_{DS} and v_{RS} depend on the optimal values for effort and consumption. $\overline{B}(\cdot)$ and $\underline{B}(\cdot)$ are negative if $P > (1+r)L$. That is, if the penalty is higher than the loan plus interest, all clients repay. Since this case does not correspond to what we observe, the remainder of this section focuses on the case where $P < (1+r)L$.

The derivation of (5.2) is based on a comparison of utility from repayment and utility from defaulting. Given the client has chosen his effort and consumption level, $U^R > U^D$ if $v_{RF} + B > v_{DF}$ in case of failure and $v_{RS} + B > v_{DS}$ in case of success. Since $v(\cdot)$ is increasing and concave, the second conditions holds whenever the first does, resulting in the first line of (5.2). That is, if it is optimal for clients to repay in case of a business failure, it is optimal to repay in case of success as well. This is because high business income (a high value for $g(\cdot)$) reduces the difference between the scheduled repayment $(1+r)L$ and the penalty P in terms of utility. The other conditions can be derived similarly.

To analyze optimal effort and consumption levels e^* and c^*, consider the first order conditions $u' = \pi \cdot v' g_1$ and $h' = u' g_2 / g_1$. Assuming that $g_{12} = 0$, they imply that clients who are successful and repay have higher effort and consumption levels than clients who default (this follows directly from comparisons of v_{RS}, v_{DS} and v_{DF}). In addition, higher non-business income V leads to higher effort and consumption levels while higher interest rates r and penalties P lead to higher effort and lower consumption. The effect of higher loan sizes depends on whether or not clients repay.

[8]Optimization proceeds as follows. Firstly, derive conditions determining when it is optimal to repay for given effort and consumption levels and business outcome. Secondly, derive the optimal effort and consumption levels for all dominant policies. Thirdly, compare the respective utility levels and choose the policy with the highest utility.

Which type of client then is more likely to repay? Firstly, the higher B, the more clients repay. Secondly, the lower $\overline{B}(\cdot)$, the more clients repay after a business failure. Thirdly, the lower $\underline{B}(\cdot)$, the more clients repay when successful. Finally, the client's income needs to be sufficiently high to enable him to repay. Let us examine each condition in turn, assuming $P < (1+r)L$.

- Higher B.

  As discussed above, $B = B(V, W, A, \delta, \zeta, \lambda, L)$ and it increases for an increasing availability of credit records, δ, for higher loan sizes L, and for a higher leverage of the business, λ. B decreases for a higher supply of loans in the market and increasing competition, ζ. The effects of income V and wealth W are unclear.

- Lower $\overline{B}(V, P, r, L) = v_{DF} - v_{RF}$.

  We can calculate the following marginal effects

$$\begin{aligned} \frac{\partial \overline{B}(\cdot)}{\partial V} &= v'_{DF} - v'_{RF} < 0 \ , & \frac{\partial \overline{B}(\cdot)}{\partial P} &= -v'_{DF} < 0 \ , \\ \frac{\partial \overline{B}(\cdot)}{\partial r} &= v'_{RF} L > 0 \ , & \frac{\partial \overline{B}(\cdot)}{\partial L} &= v'_{RF}(1+r) > 0 \ . \end{aligned} \tag{5.3}$$

  That is, we expect higher repayment rates for clients with high income and for high penalties (assuming $P < (1+r)L$). Higher interest rates and loan sizes increase $\overline{B}(\cdot)$ and lead to lower repayment rates.

- Lower $\underline{B}(V, P, r, L, A) = v_{DS} - v_{RS}$.

  We can again calculate the marginal effects.

$$\begin{aligned} \frac{\partial \underline{B}(\cdot)}{\partial V} &= (v'_{DS} - v'_{RS})(g_1 + \Omega_V) \ , & \frac{\partial \underline{B}(\cdot)}{\partial P} &= -v'_{DS} + (v'_{DS} - v'_{RS}) \cdot \Omega_P \ < \ 0 \ , \\ \frac{\partial \underline{B}(\cdot)}{\partial r} &= v'_{RS} L + (v'_{DS} - v'_{RS}) \cdot \Omega_r \ , & \frac{\partial \underline{B}(\cdot)}{\partial L} &= v'_{RS}(1+r) + (v'_{DS} - v'_{RS})(g_1 + \Omega_L) \ , \end{aligned} \tag{5.4}$$

  where $\Omega_\bullet = g_2 \frac{\partial e^*}{\partial \bullet} - g_1 \frac{\partial c^*}{\partial \bullet}$ and e^* and c^* are optimal effort and consumption levels.

  The effects of the various parameters on $\underline{B}(\cdot)$ are less clear since they partly depend on the households' optimal choices. From the discussion above we know that $\Omega_P > 0$, $\Omega_r > 0$, and $v'_{DS} - v'_{RS} < 0$ (assuming $P < (1+r)L$). As a consequence, we can determine the sign of $\frac{\partial \underline{B}(\cdot)}{\partial P}$ only which is negative. That is, higher penalties lead to more frequent repayment. The direct effects are as follows: higher interest rates increase $\underline{B}(\cdot)$ while higher non-business income V decreases $\underline{B}(\cdot)$. The direct effect of L is unclear.

- Higher incomes V and $g(V + L - c, e, A)$.

  In case of failure, clients are able to repay only if $V \geq (1 + r)L$. Higher non-business income (and wealth) thus lead to higher repayment rates while higher loan sizes and a higher leverage λ lead to lower repayment rates.

  In case of success, higher business income $g(V + L - c, e, A)$ increases the money available for repayment. Again, we can calculate marginal effects.

$$\begin{array}{rclrcl} \frac{\partial g(\cdot)}{\partial V} & = & g_1 + \Omega_V \ , & \frac{\partial g(\cdot)}{\partial P} & = & \Omega_P \ > \ 0 \\ \frac{\partial g(\cdot)}{\partial r} & = & g_1 + \Omega_r \ > \ 0 \ , & \frac{\partial g(\cdot)}{\partial L} & = & g_1 + \Omega_L \ . \end{array} \tag{5.5}$$

  While higher penalties and higher interest rates lead to higher business income, the effects of higher non-business income V and loans L remain unclear.

We can summarize the predicted effects in the following table.

	Variable	effect on repayment
V	non-business inc.	+
W	wealth	(+)
P	penalty	++
r	interest rate	−
L	loan size	(-)
λ	firm's leverage	?
δ	credit records	+
ζ	supply/competition	-

Repayment rates should be lower for loans with high interest rates and in areas with a high supply of microfinance loans or high competition among microfinance providers. Repayment rates should be higher the more widespread the availability of credit records, the higher the penalties in case of default, the higher non-business income, and for groups with less access to alternative loans (women, low income clients). We also expect higher repayment rates for clients with higher wealth and smaller loans, although the effects are less clear. The influence of leverage is unclear. On the one hand, a high leverage indicates that the client heavily depends on outside finance making the future availability of loans a crucial issue. On the other hand, it indicates a large exposure and potential repayment problems in times of crisis. We thus would expect a positive influence for successful businesses and a negative influence otherwise.[9]

There are two additional effects we have not yet mentioned. Firstly, the success probability is also determined by the economic environment. If the success probability decreases, e. g. through an economic crisis, more clients will default. Secondly, the success probability and the client's subjective valuation of timely repayment are in part driven by idiosyncratic characteristics. A poor repayment performance in one loan thus can be interpreted as an (imperfect) signal for these characteristics and we would expect a poor repayment performance for the next loan as well.

Throughout the discussion we have made no distinction between clients who pay a few days late and those who pay very late or never. While the incentives for their decisions are similar, the relative importance of some of the parameters considered varies. For the decision whether or not to pay a few days late, for example, a penalty in form of higher interest rates or the consequences for future loans from the same bank play an important role. The possibility to lose collateral, to face a court judgement, or to have a bad credit record all are more important for the decision not to pay at all.

5.4 Data

This section provides a brief description of the data set and presents details of the market environment, repayment incentives and repayment behavior. Section 5.4.1 begins with a general description of the data set and its structure. Section 5.4.2

[9] Our model assumes that clients choose their effort and consumption levels after they learn whether or not their business is successful. If they obtain this information after they decide on consumption and effort levels, our results change slightly only. There is a direct effect on effort and consumption choice. The decision to repay, however, is still determined by the relative size of $B, \overline{B}$ and $\underline{B}(\cdot)$, with a change in $\underline{B}(\cdot)$ due to changed optimal values for effort and consumption.

then analyzes the terms and conditions of the loans and their changes over time. Finally, section 5.4.3 discusses new variables which we generate to capture the effects of increasing loan supply and competition, for example.

5.4.1 General Description

The data we use for our analysis has been provided by Caja Los Andes, Bolivia. It consists of information about 76.000 clients and 28.000 rejected loan applications and covers the time from Mai 1992 to June 2000. Caja Los Andes FFP S.A. is a registered savings and loan company with its main branch in La Paz, Bolivia.[10] In December 1999, Caja Los Andes was serving 36,815 clients with outstanding loans amounting to $US 35.9 Mio.

Caja Los Andes offers individual loans only and secures the loans through chattel items such as televisions or other household items. Besides chattel items, personal guarantees are used and larger loans can be secured by mortgages as well. When a new client applies for a loan, the loan officer records the application. He visits the client's business and estimates balance sheet data if there are no obvious reasons for a rejection of the loan (these could be the age of the client, less than one year of business experience, or a bad repayment record with other banks). The loan officer then suggests whether and for which amount this loan should be approved and a committee decides. When the client later on applies for a consecutive loan, the loan officer visits again and makes an update of the balance information. Clients with a very good repayment performance eventually obtain an automatic credit line and balance information is collected irregularly. While Caja Los Andes initially gave loans to micro-enterprises only (i. e. very small enterprises), the target group has broadened in recent years. The median loan amount disbursed has increased from $US 367 in 1992 to $US 565 in January-June 2000. The data set is discussed in more detail in chapter 3. More information about microfinance in Bolivia is provided in Rhyne (2001).

[10] The legal category FFP (Fondo Financiero Privado) has been created as an institutional form for small banks in Bolivia, see Rhyne (2001, pp. 118ff.).

5.4.2 Loan Conditions, Repayment Incentives and Behavior

This section begins with a description of the various stages a loan goes through and then discusses the details of loan conditions and repayment behavior.

5.4.2.1 The Life of a Loan

After a prospective client applies for a loan, the bank decides whether or not to grant the loan and whether to grant the full amount. Typically, the approved loan is smaller than the amount applied for (with a median of 80% for new clients). Depending on the client's repayment capacity, both agree on the number and frequency of the scheduled repayments, which typically are of a fixed amount. In most cases, the first repayment begins immediately (i. e. after a week or a month).

When payments are due, 27% are made early, 46% on the date due, and 27% are late. If payments are late, clients have to pay a penalty in form of higher interest rates which increases after 30, 60, and 90 days (from 0.83% to 1.11% to 1.38% to 1.94% per month).[11] If the client has not paid after a few days, the loan officer visits the client and delivers an official letter from Caja Los Andes reminding of the outstanding payment. Since a significant part of the loan officer's salary depends on the punctuality of his clients' payments, he has high incentives to collect due payments. In extraordinary circumstances (e. g. severe illness) Caja Los Andes might grant a postponement of payments, where interest is accumulated but no penalties are applied (0.74% of all loans). In recent years, Caja Los Andes has also agreed on a rescheduling for some loans since some of its clients were not able to repay the scheduled amounts due to excessively high debt levels and repayment obligations (0.45% of all loans). Most clients, however, eventually make their overdue payments and also pay the penalties accumulated. If the payments are overdue by more than 30 days, Caja Los Andes begins to collect collateral or takes the loan to court (685 court cases since 1996, 0.25% of all loans). Loans are never written-off in the internal accounting systems and Caja Los Andes attempts to recollect the capital by all possible means.

[11] This describes the regime applied in most recent loans.

Caja Los Andes closely monitors late payments and uses an internal risk classification depending on the maximal and the average number of days a loan is overdue. If one payment is five or more days overdue, the loan is considered problematic. If it is 30 or more days overdue or if the average of all payments is ten or higher, the loan is classified in the highest risk category and the client is not granted future loans.

5.4.2.2 Development of the Terms and Conditions of Loans over Time

In the first years of operation, a typical loan required frequent regular repayments and was of a relatively short duration. Over time, however, the frequency of repayments has decreased and the duration of the loans has increased. In 1992, the largest part of the loans required weekly repayment (76%), with 23% requiring fortnightly repayment. Over time this distribution changed towards monthly repayment (72% in 2000) and an increasing fraction of irregular payment schedules which are tailored to the needs of the clients (from 8% in 1997 to 25% in 2000). The mean length of loans in days has increased from 80 in 1992 to 528 in January to June 2000. These changes cannot be explained by increasing loan sizes alone. For loans between $US 400 and $US 600, for example, 76% required weekly payment in 1992 compared to 0.6% in 2000. The mean duration of these loans has increased from 81 days to 431 days, while the mean number of repayments has been roughly constant over time at 14.8. Interest rates have decreased slightly over time. The mean monthly interest rate for loans of a size below $US 1,000 in 1992 values and denoted in inflation adjusted Bolivianos, for example, has dropped from 2.5% in 1992 to 2.22% in January to June 2000.

All loans are secured through chattel items. The median coverage ratio (value of chattel items over loan size) has increased from 200% to 420% between 1992 and 1994 and dropped to 260% since. While chattel guarantees have always been required, the use of personal guarantees has increased considerably over time. These are mainly used for larger loans and for loans that are high relative to the client's assets or combined with a relatively low value of chattel guarantees. While in 1992 personal guarantees have been used by 12% of loans above $US 1,000, they were used by 53% of these loans in 1999. This increased guarantee requirement corresponds to the findings of Villas-Boas and Schmidt-Mohr (1999) who predict an increase in guarantee requirements when competition is high.

5.4.2.3 Benefits for Clients with Good Repayment Records

Clients taking repeat loans obtain loan sizes that correspond more closely to their desired amount. The approved amount was on average 38% below the desired amount for all first approved loan applications in 1994, for example, 20% below the desired amount for second applications of these clients and only 13% below for the 5th applications of these clients.[12] Over time, the average for first applications declined from 43% in 1993 to 15% in 1999. The increase in loan sizes can also be seen from the median loan growth between two consecutive loan applications which is 42%. Growth rates are higher for smaller loans and between the first loans, lower thereafter. For example, the median loan growth rates of clients with a first loan in 1996 was 64% between the first and the second loan, 47% between the second and the third loan, and 39% between the third and the fourth loan.

An increasing number of clients benefits from preferential loans, which have been introduced in 1996. Of all clients who had their first loan in 1997, 14% of their fifth loans and 37% of their seventh loans have been preferential loans, to give a few examples.

Caja Los Andes states that it does not grant future loans to clients with high arrears ($\geq$ 30 days for at least one payment). How does this hold up in practice? From 70,455 first loans with no high arrears, 46,074 (64%) clients have obtained a second loan. From 5,322 clients with high arrears, 368 (7%) clients obtain a second loan.[13] While the fraction is considerably lower for clients with high arrears, it is still surprisingly high when taking into consideration Caja Los Andes' official policy.

[12] For these calculations the ratio is truncated to one whenever the approved amount was higher than the desired amount. Besides reduced restrictions from the side of the loan officers, a part of the decrease could also be due to learning on the side of the clients.

[13] Throughout this analysis, we exclude loans disbursed to refinance prior loans from the analysis of repeat loans.

5.4.2.4 Market Characteristics

Since the early 1990's, the supply of micro-loans and competition have increased substantially in the Bolivian microfinance market. The fraction of Caja Los Andes' clients holding simultaneous loans from other institutions has increased from 5% to 24% from 1995 to June 2000 in La Paz, for example. The portfolio per capita (working age population) of the microfinance institutions in La Paz has approximately increased from $US 10 in 1992 to $US 80 in 2000 (see also section 5.4.3). Credit records are provided by the banking supervisory authority and have become widely used in the microfinance sector. The number of entries with bad credit records has increased from 52 in 1992 to 6,945 in 1999.

Economic growth has been moderate until 1997/98. Beginning in late 1998, however, an economic crisis emerged. GDP per capita decreased by 2.2% in 1999 (source: Worldbank). The micro-enterprise sector was hit severely by this crisis. For production businesses in La Paz, for example, median annualized growth in profits between two consecutive balance observations was -0.4% in the first quarter of 2000 and -1.9% in the second quarter.[14]

5.4.2.5 Repayment

The frequency of late payments has increased considerably over time. Portfolio at risk (the percentage of outstanding capital that is at least 30 days overdue) has increased from 0.5% in 1995 to 7.3% in mid 2000. While in 1995 payments were 0.9 days late on average, they were 7.8 days late in the first half of 2000. Reacting on the rise in late payments, Caja Los Andes enforced timely repayment more strictly in 1999/2000 and the percentage of punctual repayments rose to 75% after being 64% in 1998. The fraction of payments a few (one to nine) days late decreased correspondingly from 28% to 13%. The fraction of payments ten or more days late, however, continued to increase from 8% in 1998 to 12% in mid 2000.

When we compare first loans of new clients with repeat loans, we find that repayment behavior deteriorates for repeat loans. For loans distributed in 1998, for

[14] For a more detailed account of the effects of the economic crisis on the micro-enterprise sector see Rhyne (2001).

example, 15% of all first loans had at least one payment $\geq$ 30 days overdue, 17% of all second loans, and 18% of all third loans. This increase in arrears could indicate that clients are aware of Caja Los Andes' strict repayment policy when they take their first loans, but then find that the policy is not so strict after all and relax. The deterioration is consistent with the findings by Schreiner (1999).

When a client faces unexpected harsh conditions, he can negotiate a postponement of payments with Caja Los Andes. While interest is accumulated, the client does not pay any punishment for the late payments and resumes regular repayment at the end of the postponed period. The fraction of loans postponed has decreased from 8% in 1992 to 0.5% in 1994 and increased to 1.9% in 1999. In the first half of 2000, the fraction was 0.7%.

Due to the increasing supply of micro-loans and the move of consumer credit companies into the same market, many households were lured into taking multiple large loans which frequently left them with regular repayment obligations which they could not pay out of their incomes. This over-indebtedness of many clients created severe repayment problems beginning in 1998. In addition, the economic crisis in Bolivia reduced the incomes of many clients. To acknowledge the reduced repayment capacity of many clients, the banking supervisory authority in Bolivia reduced the required provisions for rescheduled loans, making it cheaper for banks to reschedule loans (Rhyne 2001, p. 150). Caja Los Andes thereafter rescheduled some of the loans, 1354 in 1999 and 876 in the first half of 2000. This amounts to 3.7% of outstanding loans in 1999 and to 2.3% in the first half of 2000.

5.4.3 Additional Variables

To capture the repayment incentives outlined in section 5.3, we generate a number of new variables which are described in the following paragraphs.

Business Environment We use GDP data from the Instituto Nacional de Estadística (INE), La Paz, *http://www.ine.gov.bo* and calculate growth rates of real quarterly GDP compared to the same quarter in the previous year. The value of GROWTH is depicted in figure 5.1. Between 1996 and 1998, growth rates were close

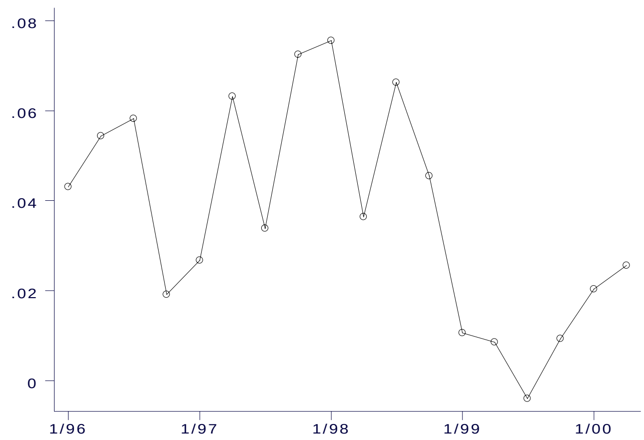

Figure 5.1: Growth in real quarterly GDP, same period (GROWTH). Source: INE

to 5% on average. In 1999 they dropped below zero and remained around 2% in 2000.

Increasing supply of loans and competition As a proxy for the increasing supply of loans we construct a variable from data about the portfolio of microfinance institutions from ASOFIN (December 1999, table 2.19) and population data from INE. From this data we calculate the portfolio per capita (working age) of the largest microfinance institutions and denote it RELPORT.[15] Its development over time is depicted in figure 5.2. There has been a strong increase over time from below $10 in 1992 to above $60 in 2000. The portfolio is highest in La Paz with a maximum of $80 in 1999.

As a proxy for competition, we calculate the fraction of clients with loans from other institutions by quarter and branch, which we denote OTHERLOAN. While this fraction increases with the supply of loans, it strongly depends on competition which can induce clients to take loans from multiple institutions at the same time in spite of additional transaction costs. The increasing number of clients with loans from other institutions is seen with concern by the microfinance institutions

[15]While this information is imperfect since it does not cover all institutions (it covers member institutions of ASOFIN, Cipame, and Finrural only), it is the only available information about supply.

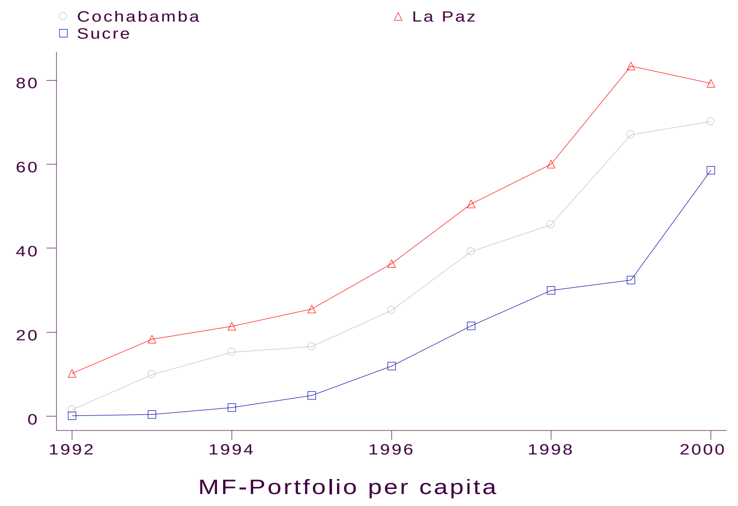

Figure 5.2: Proxy variable for the supply of microfinance loans in the branches Cochabamba, La Paz, and Sucre: RELPORT.

and seems to capture competition quite well.[16] OTHERLOAN varies considerably between the branches and it increased substantially in 1998-2000, see figure 5.3. OTHERLOAN and RELPORT are compared in figure 5.4.

Availability of credit records The data set contains information about bad credit records only. While these are high especially in the years of the economic crisis, they provide a reasonable proxy for the number of credit records as a whole. We use the number of new bad credit records each quarter as a proxy variable and denote it NEWBLOCK. It has increased from 50 per quarter in 1992 to 2000 per quarter in the first half of 2000.

Enforcement of repayment As a proxy for the fervor used to enforce timely repayment we use the ratio of payments one or two days late relative to punctual payments per quarter and denote it ENFORCE. Payments that are one or two days late most likely are caused by negligence rather than by an inability to repay and a decrease in these payments proxies the enforcement of punctual repayment. Figure 5.5 shows the development of this ratio for the three oldest branches. The

[16] An exception are the first quarters after the opening of a new branch, where the fraction of clients with other loans is very high.

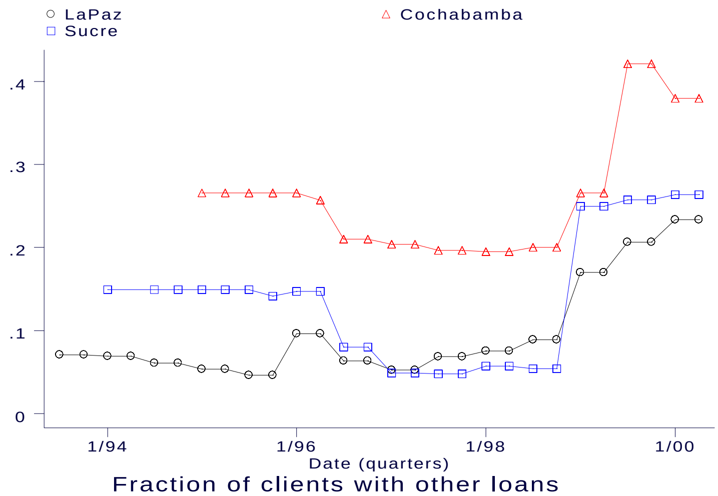

Figure 5.3: Proxy variable for competition in the branches Cochabamba, La Paz, and Sucre: OTHERLOAN.

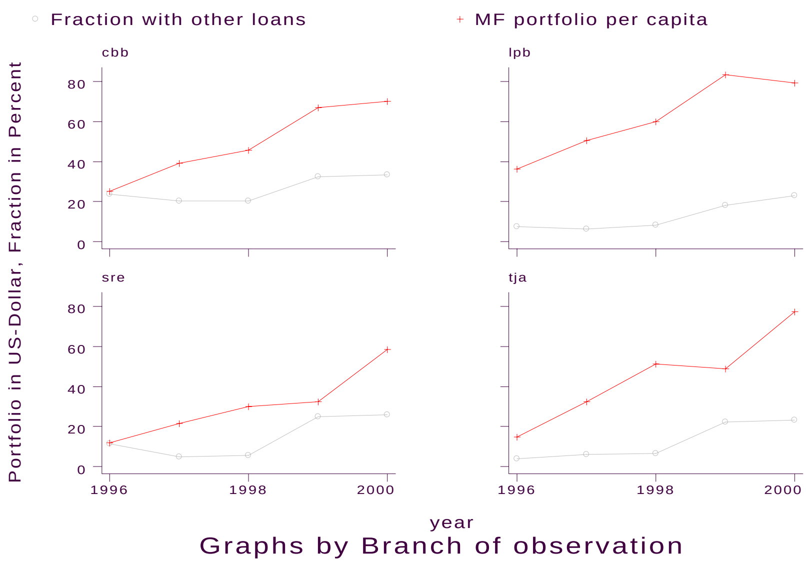

Figure 5.4: Proxy variables for the supply of microfinance loans and competition by branch (Cochabamba, La Paz, Sucre, and Tarija): OTHERLOAN and RELPORT.

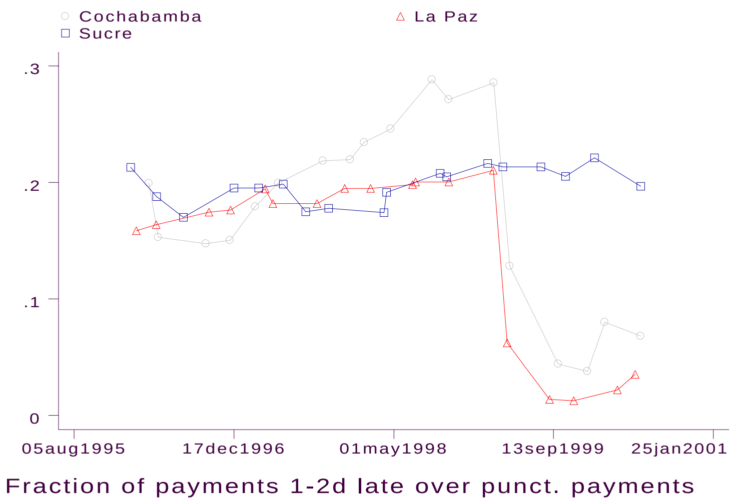

Figure 5.5: Enforcement of punctual repayment in the branches Cochabamba, La Paz, and Sucre: ENFORCE. High values indicate low enforcement, low values high enforcement.

fraction of late payments increases gradually until 1998 and drops sharply thereafter in Cochabamba and La Paz.[17]

Some of the variables described above are generated from the same data on loans that we use for our regressions and we will continue with a brief discussion of identification and endogeneity issues. Take OTHERLOAN, for example. How can we identify the effects of an increase in the fraction of clients with other loans when we also use a dummy variable for a client's individual level of indebtedness as explaining variable? Identification is possible since OTHERLOAN is an average of all clients (the fraction of clients with other loans) and thus measures a characteristic of the market rather than an individual characteristic. That is, a high value for OTHERLOAN indicates a high level of competition for all clients at the same time in the same branch, while—from a client perspective—having loans from another institution indicates that the client is more active in the loan market for a given level of competition. There are many clients who have no loans from other institutions

[17] Anecdotes about stronger enforcement abound. In La Paz, loan officers met on weekends and collectively went to their clients in order to cash in due payments. In addition, the computer system was modified to issue early warning signals if clients were overdue.

even in a market where most other clients do so (corresponding to a high value for OTHERLOAN). Identification thus is possible because we use the individual level of indebtedness and the market level of indebtedness as separate explaining variables. Moreover, OTHERLOAN is a valid independent variable since we take the mean by quarter and branch. While the mean may influence individual behavior, the influence of individual behavior on the mean is negligible due to the large number of observations and there is no endogeneity. We can make a similar case for ENFORCE which is also calculated from client data.

5.5 Estimation Strategy

The empirical analysis focuses on the prediction of loan default and late payments. The analysis uses a two-fold approach in that we use two different units of observation. Firstly, we consider loans. The analysis of loans is in the spirit of credit scoring models, predicting which loans are likely to be overdue frequently or by a long time.[18] Its results are particularly helpful for future decisions about which loans to approve and which to reject. However, we observe the full duration of these loans only if they are distributed no later than mid-1998. As a consequence, we lose valuable information since we cannot incorporate loans distributed in the years of the economic crisis and of increasing levels of indebtedness. In addition, the analysis of loans poses statistical difficulties since the loans are characterized by frequently scheduled payments and the structure of the loans (number and timing of scheduled payments, duration etc.) varies strongly. To acknowledge these issues, we change the unit of observation and supplement the analysis of loans with an explicit analysis of each payment in turn.

The remainder of this section is organized as follows. Section 5.5.1 begins with the analysis of loans and discusses estimation issues. The unit of observation then is changed in section 5.5.2, where we analyze payments. In particular, we first

[18]While most banks in developed countries use some variant of credit scoring to decide whether or not to grant a loan, this procedure is rarely used in microfinance institutions who emphasize the importance of a personal relationship between the loan officers and their clients. For more information about the applicability of credit scoring to microfinance see Schreiner (1999).

examine the probability that a payment is late and then estimate the number of days a payment is late with a duration analysis.

5.5.1 Prediction of Loan Default

For the prediction of defaults, many banks take recourse to credit scoring which traditionally is based on discriminant analysis. This paper uses a random utility model instead which allows us to estimate the probability that a loan is problematic based on individual characteristics. Discriminant analysis yields consistent estimates only if the independent variables are normally distributed. Since this assumption is violated for many of the independent variables used in our model, notably for the dummy variables, a discriminant analysis would yield inconsistent estimates (Press and Wilson 1978, McFadden 1976).

The analysis of repayment behavior is based on data from clients with approved loans. This data is no random sample of micro-entrepreneurs since we can expect rejected applicants to differ considerably from clients who obtained a loan. Further along time, clients obtain repeat loans only if they have a good repayment record. This selection can considerably bias the results of a repayment analysis since loans are approved only if the loan officer thinks that the client will repay the loan. If we predict repayment behavior of the general population from clients who obtained a loan, our estimates are biased towards good repayment behavior (assuming the loan officer's assessment is correct on average).

We can describe the selection structure of the model with a random utility model based on the theoretical model presented above. Clients repay if $U^R > U^D$ and default otherwise. Let X_0 denote the observed client characteristics that determine whether or not a client repays. X_0 thus includes loan size, interest rates, penalties, the client's income and wealth, and so on. A linearized representation of the two different utilities can be denoted as $U^R = \beta^R X_0 + \epsilon^R$, and $U^D = \beta^D X_0 + \epsilon^D$. Let the binary variable D denote default, where $D = 1$ if a client defaults and 0 otherwise. The probability that a client defaults thus can be expressed as $P(D = 1) = P(U^D > U^R) = P((\beta^D - \beta^R)X_0 + \epsilon^D - \epsilon^R > 0) = P(\beta X_0 + \epsilon_0 > 0)$. Whether or not a client defaults, however, is observed only if the client's first loan application is approved. Let z denote a binary variable with a value of one if the loan is approved and let X_1 denote client characteristics that determine this probability. Formally,

$$\begin{aligned} P(z=1) &= X_1\gamma + \epsilon_1 \quad , \quad \text{and} \\ P(D=1) &= X_0\beta + \epsilon_0 \qquad \text{with} \quad (\epsilon_1, \epsilon_0) \sim N(0,0,1,1,\rho) \quad , \end{aligned} \tag{5.6}$$

where we assume that the errors are jointly normally distributed and ρ is the coefficient of correlation between ϵ_1 and ϵ_0. The probability of default, conditioned on loan approval, then can be written as

$$P(D=1|z=1) = \frac{P(D=1 \cup z=1)}{P(z=1)} = \frac{\Phi_2(\beta X_0, \gamma X_1, \rho)}{\Phi(\gamma X_1)} \quad , \tag{5.7}$$

where Φ, Φ_2 are the univariate and bivariate normal cumulative probabilities. Selection is of no consequence if $\rho = 0$. If $\rho < 0$, clients who are more likely to obtain a loan are less likely to default and $P(D = 1|z = 1) < P(D = 1)$. Equation (5.7) can be estimated by maximum likelihood (e. g. through a bi-variate probit with selection), see Boyes, Hoffman, and Low (1989) or Greene (1998) for more details.

While the above discussion has focused on the estimates for clients with first loans, the structure for second loans is similar. The corresponding selection process describes whether or not clients obtain a second loan. That is, $z = 1$ if they obtain a second loan and $z = 0$ otherwise. In this case, ρ measures the correlation between the probability of obtaining a second loan and the probability of default. For the analysis of second loans we disregard potential selection effects stemming from the approval of the first loan for two reasons. Firstly, we find that these selection effects are insignificant even for first loans. Secondly, the results of this estimation should help the bank to decide whether or not to grant a second loan. As such, they are not meant to be applicable to the general population but to clients after a first loan.

While the two-step estimator is consistent, identification is weak if $X_1 = X_0$. If the same factors determine whether or not a client obtains a loan and whether or not he pays late, the effects of these factors are hard to pin down and identification completely relies on modeling assumptions (e. g. the joint normality of the error terms). One possibility to circumvent this problem is to impose exclusion restrictions, that is, to include variables in X_1 (selection equation) that are not included in X_0 (main equation).[19] When estimating the probability of loan approval, we include the amount applied for and its higher order terms which are excluded from the estimation of default. The amount applied for is mostly driven by a lack of personal funds or by unexpected cash shortages, both of which do not determine the client's future repayment behavior. When estimating the probability of a second loan, we use the length of the prior loan in days as explaining variable which is excluded from the estimation of defaults. The length of the loan is mainly determined by the client's repayment capacity and by the use of the loan and should thus be unrelated to future repayment behavior.

After discussing the estimation strategy, we have to define what exactly we mean with "default." To capture repayment behavior, the analysis requires a measure of default that applies to all types of loans, in particular to loans with different repayment schedules. To reduce the influence of the number of scheduled repayments, we consider the average number of days overdue per payment for each loan. That is, if a client has a loan with ten scheduled repayments and is six days late for the last two payments, the average would be 12/10=1.2 days. The distribution of this average, however, depends on the number of scheduled payments. Given an identical mean, a higher number of payments leads to a lower variance. In addition, there might be intrinsic differences in the expected number of late days between, say, weekly and monthly payments which can be captured by dummy variables.

One way to circumvent these issues would be to confine our estimates to loans with identical repayment structures. Doing this, however, would strongly bias our sample since Caja Los Andes has changed its policy of distributing loans over time. In later years, duration of loans has increased and the repayment frequencies have decreased, see section 5.4.2.2. That is, clients who would have obtained a loan with 10 weekly installments in 1995 most likely would have obtained a loan with a longer duration and, say, 5 monthly installments in 1999. Acknowledging these changes, we pool loans with different repayment structures and correct for different means and variances through the use of a robust variance estimator, through dummy variables for weekly and fortnightly schedules, and for the number of scheduled payments,

[19] See Vella (1998) for a more detailed discussion of identification issues.

and through the estimation of a heteroscedastic probit model. For the average to be a meaningful measure, we exclude all loans with less than six installments.[20]

Our analysis distinguishes between loans with few late payments that are overall unproblematic and between loans with many very late payments that increase the bank's capital at risk substantially and we calculate two sets of estimates. The first set calculates the probability that a loan has average arrears of at least one day (we denote this a "late loan" from now on) and the second calculates the probability that a loan has average arrears of at least ten days (we denote this "loan default" from now on). We choose the second threshold (ten days) since it corresponds to Caja Los Andes' internal risk evaluation. Clients with average arrears of ten days or more belong to the highest risk category and—in most cases—do not obtain future loans.

5.5.2 Analysis of Payments

While the unit of observation was the loan in the previous section, we now turn to individual payments since they allow a more direct assessment of the influence of the economic environment at the time a payment is due. Moreover, the statistical difficulties in the analysis of loans are not prevalent in the analysis of payments, the latter thus provides a test for the robustness of the results. We analyze payments in two steps. Section 5.5.2.1 begins with a probit analysis to determine payments that are at least one day late and those that are at least 10 days late. Section 5.5.2.2 then analyzes the length of each payment spell in the form of a duration analysis.

5.5.2.1 Late Payments

To begin with, we estimate the probability that a payment is at least one day late and the probability that a payment is at least ten days late. In analogy to above we refer to the former as "late payments" and to the latter as "defaults." Since there is

[20]Means and standard deviations of average arrears differ by payment form, they are 2.5, 5.05, and 6.58 (means) and 31.6, 16.8, and 32.3 (variances) for weekly, fortnightly, and monthly schedules. While the underlying distributions differ to some extent, pooling the estimates is necessary to avoid a biased sample. Control regressions using other measures than the average numbers of days overdue—e. g. the maximum number or a dummy variable indicating whether a part of the loan is unpaid after the final payment was due—show a similar structure of the results, except for the estimated influence of the payment schedule.

detailed payment information since mid 1995 only, we exclude all loans distributed before 1996.[21] For the analysis of payments, the number of the payment is very important. Shortly after the loan has been disbursed, the client should find it easy to pay the installments since he still has access to the loan. Eventually, however, the loan-money has been used for other purposes and repayment becomes more difficult. For loans with 10 scheduled repayments, for example, 1% of first payments are late, 6% of fifth payments, and 10% of final payments. In addition, prior late payments strongly increase the probability of being late for the consecutive payment.

The ideal way to take into account this time dependence would be to confine the analysis to loans with identical repayment schedules. As mentioned above, however, the policy of loan distribution has changed over time and such a sample would be biased (frequently scheduled payments, for example, were used for clients with potential repayment problems in 2000 while they were used for most clients in 1996). To reduce the influence of these changing characteristics, we again disregard loans with less than six installments and confine our analysis to the first, middle, and final payments for each loan. The analysis disregards selection effects of the form discussed in section 5.5.1 and includes first loans as well as repeat loans.

5.5.2.2 Duration Analysis

Besides the probability that a payment is late, we are interested in the number of days each payment is overdue. After how many days overdue does a repayment any time soon become unlikely? That is, if a payment is overdue for two days, for example, what is the probability that it is made the next day? How does this probability change when the payment is three days overdue? This time pattern is important for the determination of repayment policy and the decision when to call in overdue loans.

To describe this time dependence formally, let t denote the time a payment is outstanding (days outstanding = days late +1), where $f(t)$ denotes its density function, and $F(t)$ its cumulative distribution function. We then can define the

[21] Earlier loans contain information about the maximum number of days overdue only.

survivor function as

$$S(t) = P(T > t) = 1 - F(t) \quad , \tag{5.8}$$

where $S(t)$ determines the probability that the payment has not been made on day t after it was due. An alternative characterization of the time structure is given by the hazard rate, which is defined as

$$\lambda(t) = \frac{f(t)}{S(t)} \quad . \tag{5.9}$$

$\lambda(t)$ determines the probability that a client pays the next day, given that the payment is t days outstanding today. The shape of $\lambda(t)$ depends on the underlying distribution of t. To model the influence of individual characteristics on the hazard rate we use a proportional hazard rate model which specifies the hazard rate as $\lambda(t, x, \beta) = \phi(x, \beta) \cdot \lambda_0(t)$. This specification assumes that personal characteristics as captured in x proportionally shift the hazard rate $\lambda(t, x, \beta)$. That is, higher non-business income, for example, leads to a lower probability of paying on the first day overdue and to a proportionately lower probability of paying on the thirtieth day overdue. Our estimation is based on the Cox-proportional hazard rate model which assumes $\phi(x'\beta) = e^{x'\beta}$ and non-parametrically estimates β. The model allows stratification, that is, $\lambda_0(t)$ may differ among groups of observations while the influence of individual variables (β) is the same for all observations. The estimates are derived with maximum likelihood. For more details, see Kiefer (1998), for example. We use the same sample as before and again focus on first, middle, and final payments.

5.6 Estimation Results

This section presents the estimation results and is organized identically to the previous section.[22] We begin with a brief description of the selection processes relevant for

[22] For the estimation, we adjust the sample as follows. The analysis is confined to the commerce, production, and service sectors. Information from the youngest branch, Santa Cruz, is dropped since it began its operations in 1999 only. Loans and balance observations are kept only if the dates match and the calculated "income after repayment" is consistent (leading to a loss of 40% of our observations). Estimates for the second selection process (existence of a 2nd balance observations) are calculated for those clients only who had their first loan in or before 1997 to allow sufficient time for a second balance observation to occur. Observations with missing data were not used.

our analysis in section 5.6.1. Section 5.6.2 then presents the results for the analysis of loans, and section 5.6.3 changes the unit of observation and discusses payments.

5.6.1 Selection Processes

5.6.1.1 Loan Approval

The estimate of the probability of a loan approval is based on basic data available for all applications such as the amount applied for, business sectors, civil status, and branch specific information.[23] The results of the probit estimates are displayed in table 5.2. We find that loans of a size between $US 150 and $US 400 are most likely to be approved while the approval of smaller and larger loans is less likely. Having a bad repayment record with other banks (being on Caja Los Andes' black list which is based on credit records from the banking supervisory authority) has a significant negative impact on loan approval. Single clients are less likely to obtain a loan than others and clients in the commerce sector are more likely to obtain a loan (relative to the production sector).[24] The probability decreases when there are many new bad credit records (NEWBLOCK) and in areas with high competition (OTHERLOAN), while it increases in areas with a high supply of microfinance loans (RELPORT). That is, the probability that the application of a client of a given risk category is approved is higher in areas with a higher supply of loans. This finding corresponds to theoretical analyses predicting that an increased inflow of funds leads to the disbursement of loans to riskier clients (Holmström and Tirole 1997, Bolton and Freixas 2000). The change in the client structure is also evident from a comparison of new clients' characteristics over time presented in table 5.3. The mean of the log of non-business income has decreased over time, reducing the regular income

While a substantial amount of data was lost this way, the selection is not systematic since most inconsistencies are due to errors when the data was entered.

[23] There is balance information for a part of all rejected loan applications. During the time period examined here, however, the sample of rejected applicants with balance information is very small (181 observations). To asses the influence of these variables I ran separate regressions including these variables which are available on request. The estimated coefficients of the main equation are qualitatively identical.

[24] We do not control for gender since it is frequently not recorded for rejected applicants (although marital status is).

from which to repay the loan in case of business problems. Liabilities over assets and the fraction of clients with other loans, on the other hand, have increased (all these changes are significant at the 1% level). Summary statistics for approved and rejected applications can be found in table 5.4.[25] The model is significant with a Chi2 value of 8317.9 (18 degrees of freedom) and the explanatory power is relatively good with a Pseudo-R^2 of 0.33 and 84% correct in-sample predictions (using a threshold of 50%).

5.6.1.2 Observation of Second Loans

Next we estimate the probability that a client takes out a second loan given that the first loan application has been approved. We find that the observation of a second loan is more likely for women, for non-singles, younger clients, older businesses, and clients with higher incomes. The probability of a second loan diminishes strongly if the client has a bad credit record before the first loan or a bad repayment record as captured by average and maximal days overdue.[26] This dependence reflects Caja Los Andes' policy of rejecting applications from clients with a bad repayment record. The model is significant with a Chi2 value of 3686.39 (36 degrees of freedom) and the explanatory power is relatively high with a Pseudo-R^2 of 0.42. The full results of the probit estimates can be found in table 5.5, the dependence of repayment for the second loan on prior arrears is summarized in table 5.6, and additional summary statistics are listed in table 5.7.

5.6.2 Prediction of Late and Default Loans

This section discusses the results of the bivariate probit estimation for the probability that a loan is late and the probability that there is a loan default. As defined in section 5.5.1, late loans are characterized by an average of at least one day overdue per payment, while default loans are characterized by an average of at least ten days. The results for both sets of estimates can be found in table 5.8.

[25] For a more detailed discussion of this selection process see section 4.5.

[26] We include both the maximum and the average number of arrears since they capture different aspects of repayment behavior, particularly so for loans with many scheduled repayments.

The independent variables can be divided into four categories. The first are personal characteristics such as marital status, gender, age and prior repayment behavior. In addition, we include the clients' business and non-business income. These variables capture the idiosyncratic effects discussed in our theoretical model. We expect clients with less access to alternative sources of funds (clients with low income, women), with a higher risk tolerance (a higher number of previous days overdue), or with lower expected business and non-business income to pay late more frequently than others. The second category contains information about the clients' businesses such as the amount of assets, business age and business sector. We also include the ratio of liabilities over assets and a dummy variable if the client has loans from other sources which we interact with year dummies. Our model predicts a positive influence of assets on repayment while the influence of a high liabilities over asset ratio is unclear. The third category contains the terms and conditions of the loans such as interest rates and repayment schedules. The fourth category contains information about the market environment. Here we include the variables OTHERLOAN and RELPORT to capture competition and the supply of micro-loans, and the variables ENFORCE, NEWBLOCK and GROWTH (see section 5.4.3). From our model, high competition and supply should lead to worse repayment while an increased number of credit records, higher enforcement, and higher growth should lead to better repayment.

The bivariate probit estimates are presented in columns (1), (4), (5), and (6) of table 5.8. Columns (2) and (3) present slightly different sets of estimates discussed below. With respect to personal characteristics, we find that being single, young, or on the black list increases the probability of late loans and defaults. For second loans, the size of average and, in particular, maximal arrears during the first loan has a very strong negative influence on repayment behavior. As predicted, we find a high influence of risk-related, idiosyncratic characteristics.

Late and default loans are more likely for businesses with a high liabilities over asset ratio. For first loans, there is an interesting change over time in the influence of loans from other sources. Clients with prior loans from other sources were less likely than others to be late or default for loans distributed in 1996 (coefficient on "D(other loan)") but more likely than others for loans distributed in 1997 and 1998 (coefficient of the interaction terms). These findings are consistent with a prediction from our model: a higher dependence on outside funds should lead to better repayment behavior in a good economic environment (loans distributed in 1996) and to worse repayment behavior in an adverse economic environment (loans distributed in 1997 and 1998). An alternative interpretation, however, could lie in an unobserved change in the structure of other loans. While most other loans were from microfinance institutions in 1996, a larger part was from consumer credit companies in 1998. Consumer loans are more likely to exceed the clients' repayment capacity and thus could lead to more frequent late payments. Since we do not observe the source of the other loans, we cannot control for these effects.

The environment at the time the loan is distributed plays a significant role only for the probability that a loan is late, not for the probability of a default. Higher competition as approximated by OTHERLOAN leads to a higher probability, while a higher supply of loans (RELPORT) leads to a lower probability of late loans. There are significant time effects: the probability that a loan is late increases for loans distributed in 1997 and 1998, while the probability of a default is not affected. The analysis of the business environment at the time the loans are distributed, however, provides limited information for the analysis of repayment for the loan since the environment changes over the course of the loan. This effect could also explain the unexpected positive sign on the GROWTH and NEWBLOCK variables implying that higher growth and an increase in the number of bad credit records lead to worse repayment. The change in the environment over the course of the loan is explored through the analysis of payments in the next section.[27]

Terms and conditions of the loan also have a significant influence on repayment behavior. These variables have to be interpreted cautiously, however, since they suffer from endogeneity bias. That is, when a loan officer suspects that a client is not going to repay very well, he might distribute a small loan only with relatively frequent repayments and require extra guarantees. This is most evident in the ratio

[27]In principle, we could calculate statistics describing the "average" competition during the course of the loans, for example. However, this would not lead to models usable for prediction.

of applied over approved amount. Clients who obtain a loan relatively far below their desired amount are more likely to be late or default. In addition, higher loan sizes (for first loans), higher interest rates (for first loans), the existence of a personal guarantee, and fortnightly or weekly repayments (compared to monthly repayments) increase the probability that a loan is late or default. These dependencies suggest that the loan officer's assessment provides a good indicator for the client's repayment behavior. Clients who are seen to be potentially bad risks are given relatively small loans, are required to repay frequently, and a personal guarantee is required. Our results show that clients whose loans have these characteristics indeed have a worse repayment behavior than others.

To assess the robustness of our results with respect to endogeneity bias, we run separate regressions excluding variables determined by the loan officer. The results are presented in column (2) of table 5.8. While most coefficients remain qualitatively similar, the coefficients on the year dummies change and the coefficient of the clients' income become negative. That is, clients with higher income are found to pay more punctual than others as predicted by our theoretical model. This change in coefficients suggests that the loan officer determines the conditions of the loan such that they offset the negative incentive effects of low incomes. When we estimate a heteroscedastic probit (see column 3 in table 5.8) the results show a significant influence of payment schedules on the variance. While the variance equation is significant at the 1% level, the coefficients of the main equation remain qualitatively similar.

Our estimates for first and second loans show that selection effects (ρ) are insignificant. That is, we do not find evidence that Caja Los Andes selects those clients that have a good repayment behavior and we do not find evidence that clients with second loans have a better expected repayment behavior than clients without second loans. The finding that Caja Los Andes does not seem to select clients according to their repayment behavior is no contradiction to the fact that their policy is successful. On the contrary, the lacking evidence of selection bias makes our results applicable to all applicants, not only to clients.

The overall explanatory power is relatively high for late loans, but low for defaults. The out-of-sample predictive power presented in tables 5.9 and 5.10 is relatively low. Using a threshold value of 15%, for example, we correctly predict 631 late loans (40% of all late loans) but incorrectly classify 656 loans that are not late (27% of all loans that are not late). Taken together, we correctly classify 59.7% of all observations only.[28] The low predictive power, however, is common to credit scoring models of this type (Greene 1998, Schreiner 1999, for example). Our analysis of second loans fares slightly better. Using a threshold of 40%, for example, we correctly predict 545 late loans (54%) while we incorrectly classify 330 (28%). Taken together, we classify 64% of second loans correctly.

5.6.3 Analysis of Payments

After the analysis of loans, we now change the unit of observation and examine individual payments. This analysis allows us to examine the influence of the economic environment at the time each payment is due. In addition, the analysis sheds light on a change in repayment behavior between 1998 and 2000. While there has been an increase in the number of clients paying punctual, the number of clients with payments overdue for many days has increased during this period (see table 5.11).

5.6.3.1 Late Payments

This section discusses probit estimates for the probability of late and default payments, where we define a late payment as a payment that is ≥ 1 day overdue and a default payment as a payment ≥ 10 days overdue. The results are presented in table 5.13, descriptive statistics can be found in table 5.14. Columns (1) to (3) contain estimates for late payments, columns (4) to (6) contain estimates for default payments. We report results for first payments, final payments, and pooled estimates for first, middle, and final payments. Column (6) excludes potentially endogenous variables to check the robustness of the results. The pooled sample allows us to analyze the influence of the number of the payment, the number of the loans, and

[28]The threshold defines the level used for classifying loans as late. In this example, we classify a loan as late if the predicted probability to be late is at least 15%. Since our sample is unbalanced, using the typical threshold of 50% would not be adequate.

prior arrears. Since the use of multiple observations of the same clients leads to heteroscedasticity, we use robust variance estimates.[29]

The analysis of payments broadly shows the same influence factors as the analysis of loans and we discuss the additional findings only. Perhaps suprisingly, we find that women are more likely to pay late than men when we consider the full sample, so are old businesses. Being late in prior loans and prior payments of the current loan are highly significant indicators for future late and default payments. In contrast to the analysis of loans, we find no significant influence of the liabilities over asset ratio. Having loans from other sources increases the probability of late payments, especially so in the years 1999 and 2000 (see the interaction with the dummy variable for 1999/2000). The effects of other loans in 1999 and 2000 are stronger for default payments than for late payments. Taken together, this result corresponds to a prediction from our model: clients with large/multiple loans face higher repayment problems in times of the economic crisis (the year corresponds to the time the payment was due). As mentioned above, however, this effect could also be caused by an increase in consumer lending.

Our pooled estimates (columns (3), (5)-(6)) show that the probability of a late payment is considerably higher for middle and final payments and for repeat loans, while the probability of a default payment is significantly lower for repeat loans. There are two possible explanations for this different structure. Firstly, clients might learn that a limited number of days in arrears does not impede their future access to loans (Schreiner (1999) reports analogous findings). Secondly, there could be selection effects determining whether or not clients obtain repeat loans (this interpretation stands in contrast to the insignificant selection effects for second loans discussed above but could be explained by unobserved characteristics). That is, clients obtaining repeat loans are less likely to have default payments than others.

Competition (OTHERLOAN) and supply (RELPORT) have a highly significant negative effect on late payments when we consider all loans and all payments. That

[29]Additional estimates for middle payments show similar results and are available upon request. The influence of the terms and conditions of the loans differs considerably from the above analysis of loans. This is largely due to different samples since the analysis of payments additionally includes loans distributed between June 30th, 1998 and 2000.

is, a client with given characteristics has a better repayment behavior at a time and in a branch with high competition and high supply of micro-loans than elsewhere. This dependency could have two possible sources. Firstly, clients could be more aware of the importance of timely repayment in an environment with a high availability of micro-loans. Secondly, being aware of the possible negative incentive effects of high supply, institutions could have developed higher repayment incentives and/or more efficient screening to compensate for high competition and supply (through measures unobserved in our data set). Higher enforcement (a lower value of ENFORCE) leads to a lower probability of making both late and default payments (defaults to a lesser extent though). The latter effect indicates that Caja Los Andes' strengthening of enforcement has not only reached "sloppy" clients but has also decreased the probability of default payments. The effect of GROWTH is negative as well, indicating that low economic growth leads to worse repayment, although the coefficients are insignificant. Since we include dummy variables for the different years, GROWTH mainly captures variations within the years which we find to be insignificant. In two sets of control regressions firstly without year dummies and secondly with an alternative measure of growth calculated from observed micro-enterprise profits, however, GROWTH remained insignificant or had a significant positive coefficient. We are left to conclude that either both measures of growth do not capture the economic environment of micro-enterprises, or that the effect of the economic environment is very small and that the deterioration of repayment in 1999 and 2000 has not been driven by the economic crisis.

The overall explanatory power is relatively good when we consider final payments or pooled estimates with Pseudo-R^2 values between 0.14 and 0.24, lower for first payments (which are very punctual on average), and generally lower for the probability of paying late than for default payments. Out-of-sample predictions are presented in table 5.15. Using a threshold of 20%, for example, we predict 2,003 late payment correctly (64% of all late payments) and wrongly predict 3,098 payments that are not late (20%). Taken together, we correctly predict 77.1%.

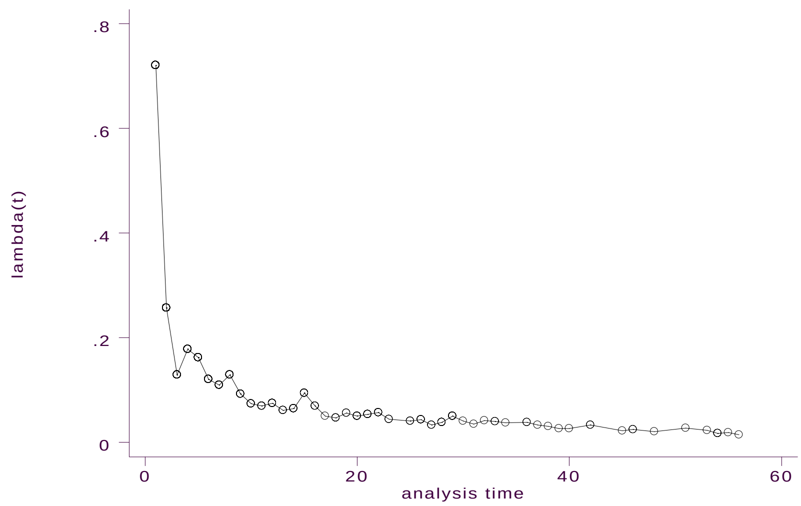

Figure 5.6: Kaplan-Meier estimates of the hazard rate function $\lambda(t)$. Interpretation: for any value of t, the value of $\lambda(t)$ indicates the probability that the client pays on day t, given the payment has not yet been made (that is, it is t-1 days overdue). Assume, for example, a payment was due on the 1st of the month and has not been made until the bank opens on the 10th of the month (it is 9 days overdue). $\lambda(10)$ then denotes the probability that the client pays on the 10th.

5.6.3.2 Duration Analysis

After discussing the probability that payments are a certain number of days late, we now examine the length of each spell a payment is (over-) due in more detail and begin with a few descriptive analyses. Figure 5.6 depicts the Kaplan-meier estimated hazard rate $\lambda(t)$ for first, middle, and final payments (see equation 5.9).[30] That is, it shows the probability that the client pays on day t, given that the payment has not yet been made (it is $t-1$ days overdue). The figure shows that the hazard rate decreases strongly initially, and slowly thereafter. That is, the probability that a client makes a due payment decreases strongly for the first few days overdue and only marginally thereafter. More than 70% of clients pay on the first day the loan is due. Of those who do not pay on the first day, only 27% pay on the second day due (that is, on the first day overdue). From then on, the fraction of clients making

[30]The estimation takes into account that a number of observations is censored since we do not observe payments made after June 30th, 2000.

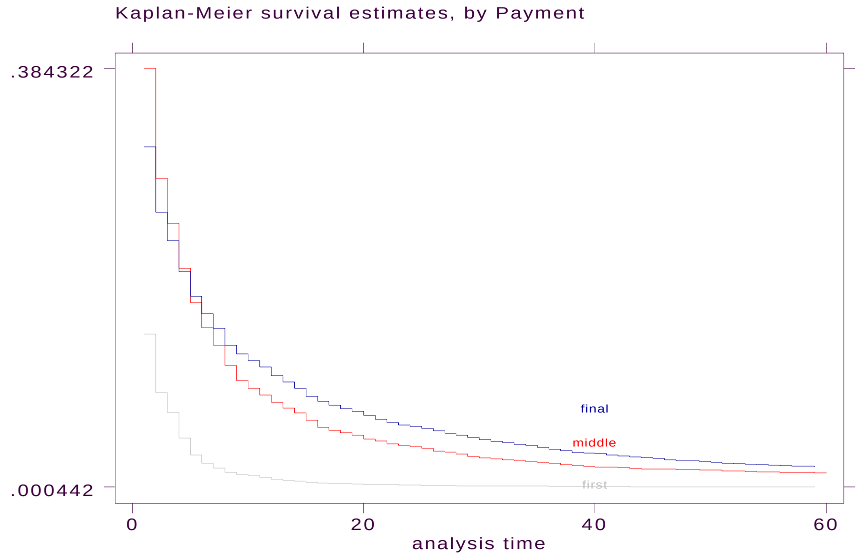

Figure 5.7: Kaplan-Meier estimates of survivor functions for first, middle, and final payments. Interpretation: for any value on the x-axis the value on the y-axis indicates the probability of being at least x days outstanding. That is, the survival estimate for 20 days, S(20), denotes the probability that the loan is outstanding at least 20 days (that is, it is at least 20-1=19 days overdue). Sample: first, middle, and final payments of loans with $\geq$ 6 installments.

their overdue payments declines to below 10%. That is, if clients are a few days late with their payments, the probability that they repay any time soon is very low. In the light of this structure, Caja Los Andes' policy to follow up on overdue payments immediately makes good sense.

To assess whether the structure of repayments differs between various groups of clients, we can plot separate estimates for these groups. For this comparison, we depict the survival functions $S(t)$ corresponding to equation (5.8). $S(t)$ denotes the probability that a payment is at least t days outstanding ($t-1$ days overdue). Figure 5.7 divides the sample between first, middle, and final payments and shows that first payments are punctual most frequently and have the lowest fraction with high arrears. The difference is highly significant with a p-value below 0.0001 (log-rank test). The mean time a payment is due is 1.59 days for first payments, 5.01 days for middle and 7.01 days for final payments. When comparing final payments in different years (figure 5.8) we find that the curves cross, indicating that repayment behavior has changed considerably. The frequency of late payments has increased from 1996 to 1998 and decreased again in 1999 (see the upper left corner of the graph). The probability of high arrears (e. g. of 40 days), however, has increased continuously over time. This graph depicts the same structure as table 5.11: while the fraction of payments that are ten or more days late has increased continuously since 1996, the fraction of punctual payments has decreased until 1998 and increased since. The differences by year again are highly significant with a p-value below 0.0001 (log-rank test) and the mean time a payment is outstanding has increased continuously from 5.12 days in 1996 to 9.78 days in 1999.

To asses the influence of prior arrears, we divide final payments into two groups dependent on whether or not the first payment of the same loan was punctual or not. Figure 5.9 shows a strong increase in the survivor function for the group with late first payments. That is, if the first payment of the loan was late, the probability that the final payment is punctual decreases from 73% to 44%, the probability that it is at least twenty days outstanding increases from 1% to 14%. The difference between the two curves again is highly significant and the mean number of days a final payment is due is 5.52 days for punctual first payments and 16.26 days for late first payments.

Besides a descriptive analysis, we can estimate a Cox-proportional hazard model as described in section 5.5.2.2 to incorporate the explanatory variables used in the other regressions. This analysis explores the full information about days overdue and not only a binary representation as the probit estimates discussed above. The estimated coefficients are presented in table 5.16 (the coefficients are in percentage terms). A negative coefficient implies that a high value of the corresponding variable decreases the hazard rate for all t. In other words, negative coefficients imply a lower probability of repayment at all times and a higher mean number of days outstanding. The structure of the results is similar to the probit estimates discussed above and we describe some prominent findings only. The probability that single clients make their payments, for example, is 1.55% lower than the probability for non-singles ($e^{-0.01557} = 0.98455$). It is 0.58% lower for women than for men and 13.69% lower

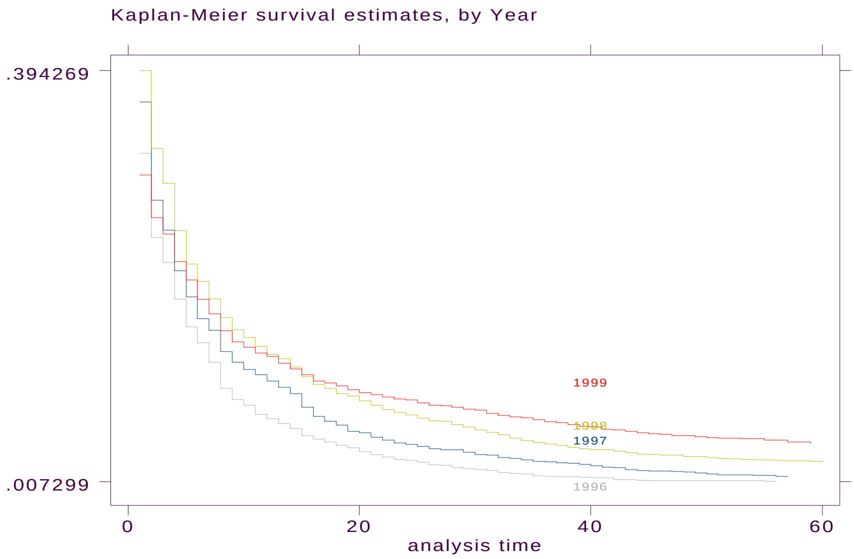

Figure 5.8: Kaplan-Meier estimates of survivor functions for final payments in 1996, 1997, 1998, and 1999. Sample: final payments of loans with ≥ 6 installments. Loans distributed since 1996.

for clients on the black list. Clients having loans from other sources have a 1.31% lower probability of making their payments. If payments for these clients are due in either 1999 or 2000, the probability decreases by a further 2.22%. Businesses in the commerce sector have a 1.23% higher probability of making their payments than businesses in the production sector.

The influence of competition as measured by OTHERLOAN is highly significant when we consider all loans and all payments and leads to an increase in hazard rates, that is, to a decrease in average time overdue. If 40% of all clients have other loans, for example, the probability that clients repay is 5.65% higher than if none of the clients had other loans ($e^{0.13745 \cdot 0.4} = 1.05652$). If only 20% of all clients have other loans, the probability increases by 2.79%. Besides the influence of competition, higher enforcement leads to significantly higher hazard rates (earlier repayment), so does a high growth rate (insignificant).

The explanatory power for first payments of first loans is very low with a Chi2 statistic of 58.80 only (44 degrees of freedom, likelihood ratio test). As mentioned

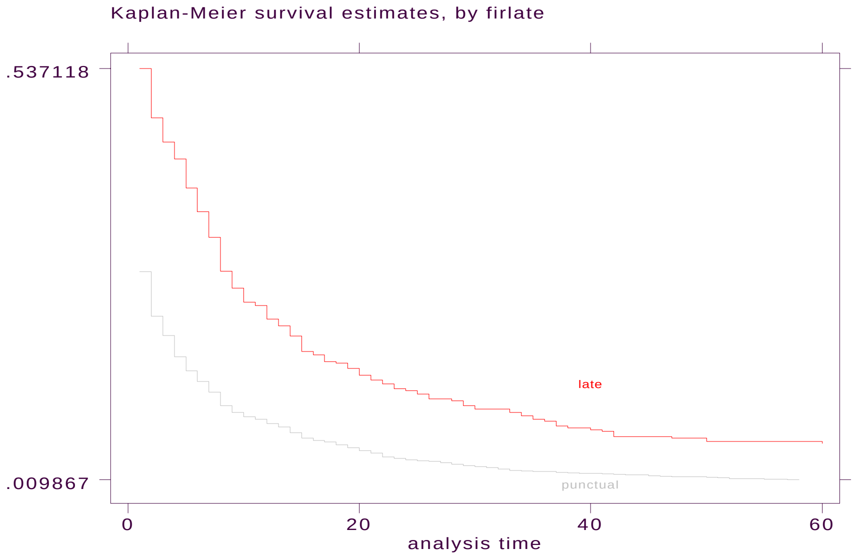

Figure 5.9: Kaplan-Meier estimates of survivor functions for final payments by punctuality of first scheduled repayment. The probability of having arrears of, e. g., 20 days is considerably higher if the first payment was late. Sample: loans with ≥ 6 installments.

above, first payments of first loans are virtually always punctual. The only significant variables are enforcement and weekly and fortnightly repayment. When we take together all payments and all loans, however, most variables are significant and the explanatory power is high with a Chi2 statistic of 6386 and 61 degrees of freedom (Wald test).[31]

The structure of the results is very similar to the probit estimates presented above. The results of all our analyses are consistent with few exceptions only and the main findings thus show a high level of robustness. With respect to our theoretical predictions, we find a strong negative influence of indebtedness in the years of the economic crisis that corresponds to our prediction. The predicted positive effect of non-business income on repayment behavior seems to be "neutralized" by the bank's repayment schedule. Surprisingly, we found no negative influence of supply and competition, and no significant influence of credit records. The examination of "penalties" suffered from endogeneity problems.

[31] Separate estimates for middle and final payments are available on request.

5.7 Discussion

The above paragraphs have discussed various aspects of repayment behavior for loans from Caja Los Andes, one of the leading microfinance institutions in Bolivia. The analysis has focused on the years 1996 to 2000 which were characterized by strongly increasing competition and supply in the microfinance market, high levels of indebtedness, and the onset of an economic crisis. How did these factors affect repayment behavior?

The fraction of clients taking loans from multiple institutions has increased substantially (from 13% in 1996 to 24% in 2000 for new clients). While there is some evidence that these clients paid more punctual than others for loans distributed in 1996, they have a worse repayment behavior than others in later years, especially so in the years of the economic crisis.

The fraction of overdue payments has increased continuously until 1998. Since then, Caja Los Andes has enforced timely repayment to reduce portfolio at risk. Our results show that increased enforcement leads not only to a higher probability that clients pay punctual, but also to a lower probability that clients have high arrears ($\geq$ 10 days).

With respect to the market environment, we find that high competition and supply, by themselves, are not responsible for high arrears. The analysis of payments shows that a client with given characteristics has a better repayment behavior at a time and in a branch with high competition and high supply of micro-loans than elsewhere. This dependency could have two possible sources. Firstly, clients could be more aware of the importance of timely repayment in an environment with a high availability of micro-loans. Secondly, being aware of the possible negative incentive effects of high supply, institutions could have developed higher repayment incentives and/or more efficient screening to compensate for high competition and supply (through measures unobserved in our data set). While the analysis of loans gives a slightly different picture with competition leading to more defaults, the analysis regards competition at the time of disbursement only.

The last two years in our sample are characterized by a severe economic crisis. Our estimates show that the influence of the economic crisis is negative but insignificant. It remains unclear whether this is because our measure of growth (changes in real quarterly GDP) does not capture the economic environment of micro-enterprises, or because the effect of the economic environment is very small and the deterioration of repayment in 1999 and 2000 has not been driven by the economic crisis (contrary to what Bolivian microlenders claim).

Throughout our analysis we find a very strong negative influence of prior late payments. If clients had overdue payments in prior loans, they are significantly more likely to pay late in future loans as well. If clients have arrears for the first payment of a loan, they are significantly more likely to have arrears for later payments as well. When analyzing the length of time a payment is overdue, we find that the probability that the client pays today decreases strongly in the first few days. That is, once the payment is three days overdue, for example, the probability that the client pays the following day is below 15% and declines further each day. These strong and highly significant influence structures suggests that an early-on focus on clients with arrears is advisable.

Taken together, our results suggest that the following factors contributed to rising arrears. Firstly, distributing more loans to clients who already have other loans leads to lower repayment rates. Secondly, clients with overdue payments in their prior loans are significantly more likely to pay late for future loans as well. This strong correlation suggests that arrears could be reduced by following a stricter policy in rejecting loans for clients with a bad repayment record stemming either from prior loans with Caja Los Andes or from the credit bureau. Thirdly, a tolerance of payments with a few days overdue leads not only to a higher probability that payments are late but also to a higher probability that they remain overdue for many days and add to capital at risk.

The analysis of the effects of terms and conditions of each loan suffered from endogeneity problems. Further research thus could use a controlled experiment to determine the influence of repayment schedules, for example. In addition, similar analyses could be carried out in other countries, notably Bangladesh, to disentangle further the effects of the economic crisis, consumer loans, and rising competition. Since all effects are concurrent in Bolivia, one could gain additional insights from an analysis in other countries where only some of these changes occur.

5.8 Appendix

Personal Characteristics	
D(single)	Marital status = single
D(female)	Gender = female
D(on black list)	Bad credit record with other banks
ln(age)	Log of the client's age
ln(non-business inc.)	Log of non-business income
ln(business income)	Log of business income
Previous average arrears	Average arrears (days overdue) in previous loan
Previous maximum arrears	Maximal arrears in previous loan
Business Characteristics	
ln(assets)	Log assets
Liabilities/assets	Ratio of liabilities over assets
D(other loans)	Dummy: client has loans from other sources
ln(Business Age)	Log business age
D(Commerce)	Dummy: Commerce Sector (relative to production sector)
D(Service)	Dummy: Service Sector (relative to production sector)
Loan Characteristics	
ln(approved amount)	Log approved amount (loan size)
Appl./appr. amount	Amount applied for over approved amount
D(preferential)	Dummy: preferential loans (automatic credit line)
Interest Rate	Interest rate
Flat rate commission	Commission charged on each payment
D(old penalty code)	Dummy: "old" penalties system for late payments
D(weekly)	Dummy: weekly repayments
D(fortnightly)	Dummy: fortnightly repayments
D(irregular)	Dummy: irregular repayments (since 1997)
Length of loan (days)	Duration of the loan in days
D(... installments)	Dummy: number of scheduled repayments
ln(value of chattel g.)	Log value of chattel guarantees
D(pers. guarantee)	Dummy: existence of a co-signer/guarantor
Environment	
D(Cochabamba)	Dummy: loan disbursed in Cochabamba (relative to La Paz)
D(Sucre)	Dummy: loan disbursed in Sucre (relative to La Paz)
D(Trinidad)	Dummy: loan disbursed in Trinidad (relative to La Paz)
D(Tarija)	Dummy: loan disbursed in Tarija (relative to La Paz)
OTHERLOAN	Fraction of clients with loans from other sources
RELPORT	Portfolio of MF institutions per capita
ENFORCE	Tolerance of one or two days overdue
NEWBLOCK	Number of new entries with bad credit records
GROWTH	Quarterly growth rate (source: INE)
D(199x)	Dummy: Year=199x

Table 5.1: List of variables used for the empirical analysis. All logs are calculated as log(<variable>+1).

	Dependent variable: approval of first loan application
ln(applied amount)	2.055
	(7.02)**
ln(applied amount)^2	-0.271
	(5.72)**
ln(applied amount)^3	0.010
	(4.16)**
On black list	-0.740
	(6.88)**
D(single)	-1.372
	(72.50)**
D(1997)	-0.065
	(0.76)
D(1998)	-0.303
	(2.25)*
D(Cochabamba)	-0.451
	(3.67)**
D(Sucre)	-0.236
	(1.62)
D(Trinidad)	0.082
	(0.30)
D(Tarija)	-0.543
	(5.49)**
D(Commerce)	0.076
	(3.28)**
D(Service)	-0.019
	(0.67)
OTHERLOAN	-2.768
	(5.05)**
RELPORT	0.028
	(5.42)**
ENFORCE	-1.328
	(2.28)*
NEWBLOCK	-6.29e-05
	(4.26)**
GROWTH	-1.743
	(2.58)**
Constant	-3.539
	(5.63)**
Observations	29356
Percentage approved	72.74
Pseudo-R^2	0.33
Chi2 (18)	8317.9
Log Likelihood	-11465.96

Table 5.2: Selection estimates for the approval of the first loan application. Robust standard errors are in parentheses, ** denotes significant at the 1% level, * at the 5% level. Sample: new applications between Jan. 1st 1996 and June 30th 1998.

	Statistics for new clients.					
	1996		1998		2000	
Personal Characteristics						
D(single)	0.208	*(0.406)*	0.192	*(0.394)*	0.285	*(0.452)*
D(female)	0.631	*(0.483)*	0.589	*(0.492)*	0.575	*(0.494)*
D(on black list)	4.38e-03	*(0.066)*	4.56e-03	*(0.067)*	5.35e-03	*(0.073)*
ln(age)	3.594	*(0.284)*	3.568	*(0.282)*	3.581	*(0.310)*
ln(non-business income)	2.671	*(2.247)*	2.597	*(2.322)*	2.448	*(2.310)*
ln(business income)	4.925	*(0.779)*	5.022	*(0.805)*	4.775	*(0.921)*
Business Characteristics						
ln(assets)	6.879	*(1.172)*	7.098	*(1.155)*	7.214	*(1.179)*
Liabilities over assets	0.018	*(0.062)*	0.026	*(0.083)*	0.057	*(0.136)*
D(other loan)	0.134	*(0.340)*	0.146	*(0.353)*	0.236	*(0.425)*
ln(Business Age)	1.491	*(0.984)*	1.261	*(0.797)*	1.804	*(0.704)*
D(Commerce)	0.557	*(0.497)*	0.522	*(0.500)*	0.486	*(0.500)*
D(Service)	0.155	*(0.362)*	0.199	*(0.399)*	0.227	*(0.419)*
Observations	6,341		4,821		1884	

Table 5.3: Summary statistics for new clients by years. Sample: A 70% random sample of first loans in or after 1996 with at least six installments.

	Approved loans N=21,353		Rejected Applications N=8,010	
	Mean	Std. Dev.	Mean	Std. Dev.
ln(applied amount)	5.788	*(1.025)*	5.869	*(1.160)*
On black list	4.50e-03	*(0.067)*	0.011	*(0.102)*
D(single)	0.195	*(0.396)*	0.743	*(0.437)*
D(1997)	0.365	*(0.481)*	0.329	*(0.470)*
D(1998)	0.178	*(0.382)*	0.181	*(0.385)*
D(Cochabamba)	0.126	*(0.332)*	0.323	*(0.468)*
D(Sucre)	0.102	*(0.302)*	0.203	*(0.402)*
D(Trinidad)	0.044	*(0.204)*	0.113	*(0.317)*
D(Tarija)	0.094	*(0.291)*	0.179	*(0.383)*
D(Commerce)	0.539	*(0.498)*	0.519	*(0.500)*
D(Service)	0.180	*(0.384)*	0.249	*(0.433)*
OTHERLOAN	0.096	*(0.061)*	0.138	*(0.081)*
RELPORT	38.06	*(14.07)*	28.17	*(11.30)*
ENFORCE	0.176	*(0.024)*	0.169	*(0.034)*
NEWBLOCK	700.0	*(807.2)*	732.6	*(844.9)*
GROWTH	0.048	*(0.018)*	0.047	*(0.017)*

Table 5.4: Mean and standard deviations by loan approval. Sample: new applications between Jan. 1st 1996 and June 30th 1998.

	Dependent variable: existence of a second approved loan		
ln(appr.amount)	0.034	ln(income)	0.153
	(1.04)		(4.42)**
D(on black list)	-0.470	Appl./appr.am.	-0.030
	(2.50)*		(2.61)**
D(single)	-0.067	Av.arrears	-0.262
	(2.03)*		(14.66)**
D(female)	0.116	Av.arrears^2	3.31e-03
	(4.21)**		(9.29)**
D(1997)	0.105	Max.arrears	-0.069
	(0.59)		(11.16)**
D(1998)	-1.507	Max.arrears^2	1.81e-03
	(5.13)**		(12.95)**
D(Cochabamba)	-0.516	Max.arrears^3	-1.14e-05
	(2.36)*		(10.29)**
D(Sucre)	0.296	Max.arrears^4	6.01e-09
	(0.90)		(8.97)**
D(Trinidad)	-1.123	Loan length	0.015
	(2.01)*		(17.32)**
D(Tarija)	0.038	Loan length^2	-5.07e-05
	(0.16)		(17.68)**
D(Commerce)	0.079	Loan length^3	3.49e-08
	(2.58)**		(12.59)**
D(Service)	0.015	D(8-10 inst.)	-0.022
	(0.38)		(0.62)
OTHERLOAN	2.836	D(11-14 inst.)	-0.023
	(3.55)**		(0.57)
RELPORT	3.33e-03	D(15-18 inst.)	-0.048
	(0.27)		(0.98)
ENFORCE	-6.035	D(19-24 inst.)	-0.069
	(5.81)**		(1.00)
GROWTH	0.416	D(25-30 inst.)	0.019
	(0.56)		(0.11)
ln(assets)	-4.19e-03	D(>30 inst.)	-0.308
	(0.20)		(1.36)
Liab./assets	0.105	ln(age)	-0.152
	(0.55)		(3.24)**
Constant	0.375	ln(bus. age)	0.039
	(0.76)		(2.50)*
Observations	17537		
Pct. with 2nd loans	66.04		
Pseudo-R^2	0.42		
Chi2 (36)	3868.39		
Log-Likelihood	-6507.86		

Table 5.5: Selection estimates for the existence of second loans. Robust standard errors are in parentheses, ** denotes significant at the 1% level, * at the 5% level. Sample: first loans distributed between 1995 to mid 1998. Variables refer to observations made when the first loan was distributed.

Arrears for second loans	Previous average arrears		Previous max. arrears		Observations
Av. Arrears < 1	0.26	*(0.80)*	1.49	*(3.93)*	12,135
Av. Arrears ≥ 1	0.70	*(1.50)*	3.26	*(5.83)*	5,610
Av. Arrears < 10	0.36	*(1.00)*	1.88	*(4.49)*	16,648
Av. Arrears ≥ 10	1.04	*(1.88)*	4.59	*(6.70)*	1,097
no 2nd loan	22.4	*(78.4)*	60.8	*(169.5)*	5,963

Table 5.6: Average arrears of current loans depending on previous repayment behavior. Standard deviations are in parentheses.

	No 2nd loan N=5,963		With 2nd loan N=17,089	
	Mean	Std. Dev.	Mean	Std. Dev.
ln(appr.amount)	5.507	*(0.997)*	5.181	*(0.907)*
D(on black list)	6.35e-03	*(0.079)*	2.28e-03	*(0.048)*
D(single)	0.226	*(0.418)*	0.184	*(0.387)*
D(female)	0.568	*(0.495)*	0.640	*(0.480)*
D(1997)	0.369	*(0.483)*	0.251	*(0.433)*
D(1998)	0.249	*(0.432)*	0.009	*(0.092)*
Cochabamba	0.147	*(0.354)*	0.138	*(0.344)*
Sucre	0.105	*(0.307)*	0.130	*(0.337)*
Trinidad	0.041	*(0.198)*	0.016	*(0.127)*
Tarija	0.104	*(0.305)*	0.095	*(0.294)*
D(Commerce)	0.485	*(0.500)*	0.580	*(0.494)*
D(Service)	0.218	*(0.413)*	0.155	*(0.362)*
OTHERLOAN	0.098	*(0.062)*	0.095	*(0.068)*
RELPORT	39.40	*(14.53)*	29.63	*(13.84)*
ENFORCE	0.179	*(0.025)*	0.173	*(0.022)*
GROWTH	0.048	*(0.019)*	0.044	*(0.022)*
ln(assets)	7.068	*(1.211)*	6.914	*(1.219)*
Liab./assets	0.022	*(0.074)*	0.020	*(0.066)*
ln(income)	5.397	*(0.574)*	5.385	*(0.573)*
Appl./appr.am.	1.615	*(1.046)*	1.662	*(1.117)*
Av.arrears	22.37	*(78.35)*	0.403	*(1.092)*
Max.arrears	60.81	*(169.5)*	2.053	*(4.703)*
Loan length	226.4	*(125.9)*	140.4	*(68.69)*
D(8-10 inst.)	0.342	*(0.474)*	0.316	*(0.465)*
D(11-14 inst.)	0.314	*(0.464)*	0.268	*(0.443)*
D(15-18 inst.)	0.113	*(0.317)*	0.096	*(0.295)*
D(19-24 inst.)	0.055	*(0.229)*	0.025	*(0.156)*
D(25-30 inst.)	0.010	*(0.100)*	3.59e-03	*(0.060)*
D(>30 inst.)	4.34e-03	*(0.066)*	1.86e-03	*(0.043)*
ln(age)	3.571	*(0.286)*	3.591	*(0.281)*
ln(bus. age)	1.391	*(0.893)*	1.546	*(0.916)*

Table 5.7: Summary statistics by existence of a second loan. Variables refer to observations made when the first loan was distributed. Sample: first loans distributed between 1995 and mid-1998.

	First loans: Dependent variable:				Second loans: Dependent variable:	
	$P(late)$	$P(late)$	$P(late)$	$P(default)$	$P(late)$	$P(default)$
	(1)	(2)	(3)	(4)	(5)	(6)
Personal Characteristics						
D(single)	0.101	0.160	0.106	0.221	0.052	-0.010
	(1.62)	(2.53)*	(3.95)**	(2.32)*	(1.51)	(0.19)
D(female)	6.04e-03	0.013	0.012	-0.012	0.039	4.46e-04
	(0.29)	(0.65)	(0.60)	(0.41)	(1.36)	(0.01)
D(on black list)	0.341	0.326	0.357	0.413	0.551	0.622
	(2.49)*	(2.41)*	(2.78)**	(2.34)*	(2.88)**	(2.74)**
ln(age)	-0.131	-0.142	-0.129	-0.086	-0.291	-0.304
	(3.63)**	(4.06)**	(3.49)**	(1.63)	(5.81)**	(4.23)**
ln(non-business inc.)	0.012	-0.006	0.011	0.004	0.019	9.44e-03
	(2.25)*	(1.22)	(2.20)*	(0.55)	(2.67)**	(0.93)
ln(business income)	0.040	-0.057	0.045	0.024	0.092	7.72e-03
	(1.88)	(2.83)**	(2.18)*	(0.72)	(2.96)**	(0.17)
Previous average arrears					0.156	0.167
					(4.12)**	(3.74)**
$(\text{Prev.av.arrears})^2$					-1.13e-03	-0.010
					(0.37)	(2.87)**
Previous max. arrears					0.157	0.132
					(10.47)**	(6.60)**
$(\text{Prev.max.arrears})^2$					-0.012	-0.013
					(8.17)**	(5.62)**
$(\text{Prev.max.arrears})^3$					2.94e-04	4.63e-04
					(5.83)**	(4.78)**
$(\text{Prev.max.arrears})^4$					-2.10e-06	-4.90e-06
					(4.63)**	(4.27)**
Business Characteristics						
ln(assets)	-0.025	-0.065	-0.019	-0.053	-0.022	-0.041
	(1.53)	(3.99)**	(1.18)	(2.13)*	(0.91)	(1.16)
Liabilities/assets	0.381	0.402	0.356	0.533	0.105	0.639
	(2.20)*	(2.36)*	(2.05)*	(2.23)*	(0.44)	(2.09)*
D(other loans)	-0.117	-0.129	-0.123	-0.108	0.162	-0.165
	(1.90)	(2.12)*	(2.17)*	(1.12)	(1.79)	(1.05)
D(other loans)*D(1997)	0.150	0.137	0.153	0.078	-0.030	0.152
	(1.96)*	(1.81)	(2.06)*	(0.68)	(0.28)	(0.86)
D(other loans)*D(1998)	0.162	0.163	0.184	0.126	-0.054	0.059
	(1.84)	(1.87)	(2.07)*	(1.02)	(0.42)	(0.30)
ln(Business Age)	0.018	0.019	0.018	-0.027	0.032	-0.013
	(1.45)	(1.57)	(1.54)	(1.47)	(1.83)	(0.46)
D(Commerce)	-0.105	-0.124	-0.098	-0.140	0.008	0.037
	(3.97)**	(5.46)**	(3.62)**	(3.72)**	(0.24)	(0.75)
D(Service)	0.038	0.062	0.045	-0.012	0.130	0.197
	(1.23)	(2.19)*	(1.47)	(0.27)	(2.97)**	(3.25)**

Table 5.8: Results from bivariate probit analysis.

Loan Characteristics						
ln(approved amount)	0.338	0.647	0.356	0.708	-0.043	0.062
	(3.45)**	(7.15)**	(3.48)**	(4.62)**	(0.30)	(0.30)
ln(approved amount)^2	-0.035	-0.042	-0.038	-0.056	-0.005	-0.007
	(4.20)**	(5.53)**	(4.19)**	(4.31)**	(0.40)	(0.40)
Appl./appr. amount	0.036		0.033	0.066	0.095	0.199
	(3.68)**		(3.43)**	(5.52)**	(2.88)**	(4.71)**
Interest Rate	0.037		0.041	0.126	0.055	0.003
	(1.08)		(1.16)	(2.59)**	(1.15)	(0.05)
Flat rate commission	-0.126		-0.144	-0.220	-0.278	0.334
	(1.31)		(1.38)	(1.78)	(1.80)	(1.46)
D(old penalty code)	0.007		0.013	-0.123	-0.091	-0.273
	(0.16)		(0.28)	(1.81)	(1.61)	(3.20)**
D(weekly)	0.391		0.443	0.041	0.226	-0.170
	(3.66)**		(3.31)**	(0.23)	(1.43)	(0.59)
D(fortnightly)	0.297		0.434	0.062	0.242	0.228
	(4.86)**		(6.12)**	(0.64)	(2.78)**	(1.57)
Length of loan (days)	0.011		9.38e-03	6.96e-03	7.12e-03	5.31e-03
	(12.78)**		(8.30)**	(5.41)**	(6.40)**	(2.86)**
Length of loan^2	-1.64e-05		-1.34e-05	-9.02e-06	-8.90e-06	-4.38e-06
	(8.38)**		(5.42)**	(3.48)**	(3.68)**	(1.29)
Length of loan^3	8.36e-09		6.67e-09	3.93e-09	4.32e-09	1.18e-09
	(5.30)**		(3.45)**	(2.08)*	(2.41)*	(0.53)
D(8-10 installments)	-0.036		-0.043	-0.024	-0.113	-0.139
	(0.93)		(1.12)	(0.39)	(1.56)	(1.28)
D(11-14 installments)	-0.109		-0.138	-0.070	-0.152	-0.245
	(1.93)		(2.31)*	(0.77)	(1.65)	(1.68)
D(15-18 installments)	-0.172		-0.243	-0.048	-0.233	-0.417
	(2.23)*		(2.81)**	(0.38)	(1.97)*	(2.17)*
D(19-24 installments)	-0.171		-0.281	-0.075	-0.303	-0.447
	(1.73)		(2.56)*	(0.47)	(2.10)*	(1.88)
D(25-30 installments)	-0.084		-0.197	0.116	-0.427	-0.643
	(0.57)		(1.23)	(0.53)	(2.43)*	(2.15)*
D(>30 installments)	-0.541		-0.700	-0.898	-0.275	-0.489
	(2.76)**		(2.63)**	(2.28)*	(1.21)	(1.19)
ln(value of chattel g.)	-0.010		-0.011	0.021	-0.009	0.025
	(2.25)*		(2.35)*	(3.21)**	(1.56)	(3.00)**
D(pers. Guarantee)	0.199		0.187	0.231	0.160	0.182
	(9.96)**		(7.38)**	(7.79)**	(5.94)**	(4.68)**

Table 8 continued.

Environment						
Cochabamba	-0.457	-0.177	-0.473	0.343	-0.048	0.564
	(3.17)**	(1.27)	(3.14)**	(1.60)	(0.21)	(1.53)
Sucre	-0.407	0.032	-0.417	-0.444	-0.201	-0.206
	(2.34)*	(0.19)	(2.17)*	(1.93)	(0.76)	(0.60)
Trinidad	-1.836	-0.979	-1.866	-0.572	-0.910	0.473
	(5.73)**	(3.18)**	(4.98)**	(1.29)	(1.85)	(0.69)
Tarija	-0.047	0.174	-0.052	0.069	-0.211	0.061
	(0.41)	(1.55)	(0.41)	(0.46)	(1.28)	(0.29)
OTHERLOAN	4.808	2.994	4.718	-0.105	2.283	-1.904
	(7.16)**	(4.72)**	(6.34)**	(0.09)	(2.00)*	(0.91)
RELPORT	-0.021	-8.37e-03	-0.022	-4.00e-03	-0.011	7.68e-03
	(3.34)**	(1.35)	(3.09)**	(0.48)	(1.13)	(0.64)
ENFORCE	-0.865	-0.812	-0.977	0.033	-1.923	-0.936
	(1.10)	(1.07)	(1.29)	(0.03)	(1.64)	(0.55)
NEWBLOCK	2.98e-05	5.86e-05	3.26e-05	-2.60e-06	-7.40e-06	-2.57e-05
	(1.81)	(3.87)**	(2.04)*	(0.11)	(0.37)	(0.89)
GROWTH	2.266	3.222	2.314	-0.074	-0.667	-2.572
	(3.20)**	(4.87)**	(3.35)**	(0.07)	(0.75)	(1.97)*
D(1997)	0.407	0.378	0.425	-0.014	0.468	-0.170
	(4.13)**	(3.95)**	(3.80)**	(0.10)	(3.29)**	(0.90)
D(1998)	0.675	0.626	0.699	0.136	0.571	-0.087
	(4.31)**	(4.08)**	(3.92)**	(0.64)	(2.49)*	(0.29)
ρ_1 (Appr. of 1st loan)	0.028	-0.043		-0.077		
	(0.28)	(0.43)		(0.52)		
ρ_2 (Exist. of 2nd loan)					0.039	0.115
					(0.44)	(0.92)
Constant	-2.282	-1.749	-2.195	-4.378	-0.557	-2.003
	(5.24)**	(4.57)**	(4.96)**	(6.74)**	(0.86)	(2.11)*
Variance equation						
Number of installments			0.013			
			(2.03)*			
D(weekly)			-0.211			
			(1.66)			
D(fortnightly)			-0.321			
			(3.65)**			
Wald Chi(3)			16.17			
Observations	29220	29221	21218	29220	17342	17342
Percentage late/default	30.0	30.0	30.0	7.28	33.33	7.16
Wald Chi(43(27)/49)	1339.02	879.31	312.16	748.57	995.89	425.06
Log Likelihood	-23556.36	-23884.65	-12110.12	-16477.34	-12982.81	-9178.48

Table 8 continued: Results from bivariate probit analysis. First loans include 8,002 censored observations, second loans 5,956. Columns (1), (4), (5), and (6) are bivariate probit analyses with selection. Column (2) excludes variables that are potentially endogenous. Column (3) is based on a heteroscedastic probit. Robust standard errors are in parentheses, ** denotes significant at the 1% level, * at the 5% level. Sample: Loans distributed between Jan. 1st 1996 and June 30th 1998 with corresponding balance observations and at least six scheduled repayments. Second loans of clients who had first loans before 1995 are excluded due to incomplete information on prior arrears.

late=1	p10 0	1	p15 0	1	p20 0	1	Total
0	1,045	1,355	1,744	656	2,148	252	2,400
	(26.2)	(34.0)	(43.8)	(16.5)	(53.9)	(6.3)	(60.3)
1	452	1,130	951	631	1,347	235	1,582
	(11.4)	(28.4)	(23.9)	(15.9)	(33.8)	(5.9)	(39.7)
Total	1,497	2,485	2,695	1,287	3,495	487	3,982
	(37.6)	(62.4)	(67.7)	(32.3)	(87.8)	(12.2)	(100.0)

Table 5.9: Out-of-sample predictive power for first loans. Calculations are based on a probit estimate for the probability that loans are late using loan data from 1996 and 1997, predictions are made for the first half of 1998. Percentages are in parentheses.

late=1	p30 0	1	p40 0	1	p50 0	1	p60 0	1	Total
0	500	674	844	330	1,010	164	1,100	74	1,174
	(0.23)	(0.31)	(0.39)	(0.15)	(0.46)	(0.08)	(0.50)	(0.03)	(0.54)
1	204	808	467	545	665	347	829	183	1,012
	(0.09)	(0.37)	(0.21)	(0.25)	(0.30)	(0.16)	(0.38)	(0.08)	(0.46)
Total	704	1,482	1,311	875	1,675	511	1,929	257	2,186
	(0.32)	(0.68)	(0.60)	(0.40)	(0.77)	(0.23)	(0.88)	(0.12)	(1.00)

Table 5.10: Out-of-sample predictive power for second loans. Calculations are based on a probit estimate for the probability that loans are late using loan data from 1996 and 1997, predictions are made for the first half of 1998. Percentages are in parentheses.

Year	0	1-9	10-29	≥30
1996	75.1	21.8	2.5	0.5
1997	68.0	25.7	4.9	1.4
1998	63.9	27.6	6.2	2.3
1999	75.5	15.5	5.4	3.6
2000	75.5	12.9	6.5	5.2

Table 5.11: Fraction of payments that are punctual, one to nine, ten to 29, or 30 and more days late (in %). Sample: all payments due between January 1st 1996 and June 30th 2000.

	late loans		not late		default		no default	
Personal Characteristics								
D(single)	0.182	(0.386)	0.224	(0.417)	0.190	(0.392)	0.264	(0.441)
D(female)	0.619	(0.486)	0.596	(0.491)	0.617	(0.486)	0.558	(0.497)
D(on black list)	3.94e-03	(0.063)	5.79e-03	(0.076)	4.24e-03	(0.065)	7.74e-03	(0.088)
ln(age)	3.587	(0.283)	3.561	(0.281)	3.582	(0.283)	3.541	(0.279)
ln(non-business inc.)	2.682	(2.278)	2.527	(2.314)	2.657	(2.284)	2.369	(2.341)
ln(business income)	4.974	(0.801)	5.037	(0.784)	4.981	(0.798)	5.142	(0.766)
Business Characteristics								
ln(assets)	6.981	(1.217)	7.122	(1.189)	7.002	(1.215)	7.286	(1.130)
Liabilities / assets	0.019	(0.067)	0.023	(0.075)	0.020	(0.068)	0.027	(0.082)
D(other loans)	0.079	(0.270)	0.093	(0.290)	0.082	(0.274)	0.098	(0.297)
D(other loans)*D(1997)	0.029	(0.167)	0.038	(0.191)	0.031	(0.173)	0.039	(0.194)
D(other loans)*D(1998)	0.015	(0.120)	0.027	(0.161)	0.017	(0.128)	0.037	(0.188)
ln(Business Age)	1.401	(0.912)	1.444	(0.861)	1.413	(0.901)	1.430	(0.850)
D(Commerce)	0.556	(0.497)	0.501	(0.500)	0.545	(0.498)	0.467	(0.499)
D(Service)	0.166	(0.372)	0.213	(0.410)	0.177	(0.382)	0.220	(0.414)
Loan Characteristics								
ln(approved amount)	5.390	(0.991)	5.579	(0.958)	5.420	(0.984)	5.791	(0.941)
Appl. / appr. amount	1.564	(0.989)	1.548	(0.958)	1.556	(0.982)	1.603	(0.949)
Interest Rate	3.214	(0.376)	3.181	(0.397)	3.210	(0.380)	3.134	(0.418)
Flat rate commission	0.996	(0.089)	0.993	(0.109)	0.996	(0.093)	0.988	(0.116)
D(old penalty code)	0.517	(0.500)	0.401	(0.490)	0.496	(0.500)	0.310	(0.463)
D(weekly)	0.162	(0.368)	0.094	(0.292)	0.148	(0.355)	0.056	(0.230)
D(fortnightly)	0.487	(0.500)	0.411	(0.492)	0.473	(0.499)	0.357	(0.479)
Length of loan (days)	191.2	(105.8)	235.8	(114.3)	200.0	(108.3)	262.4	(118.5)
D(8-10 installments)	0.373	(0.484)	0.368	(0.482)	0.374	(0.484)	0.346	(0.476)
D(11-14 installments)	0.302	(0.459)	0.327	(0.469)	0.307	(0.461)	0.343	(0.475)
D(15-18 installments)	0.108	(0.310)	0.110	(0.313)	0.108	(0.310)	0.121	(0.326)
D(19-24 installments)	0.041	(0.199)	0.055	(0.228)	0.044	(0.205)	0.060	(0.237)
D(25-30 installments)	6.21e-03	(0.079)	0.011	(0.105)	7.12e-03	(0.084)	0.015	(0.121)
D(>30 installments)	4.08e-03	(0.064)	4.54e-03	(0.067)	4.44e-03	(0.067)	1.29e-03	(0.036)
ln(val. chattel items)	5.738	(2.244)	5.661	(2.454)	5.704	(2.300)	5.855	(2.423)
D(pers. Guarantee)	0.431	(0.495)	0.527	(0.499)	0.448	(0.497)	0.612	(0.487)
Environment								
Cochabamba	0.113	(0.317)	0.156	(0.363)	0.123	(0.328)	0.165	(0.371)
Sucre	0.094	(0.291)	0.120	(0.325)	0.105	(0.306)	0.064	(0.245)
Trinidad	0.042	(0.200)	0.048	(0.214)	0.044	(0.206)	0.034	(0.182)
Tarija	0.085	(0.278)	0.115	(0.319)	0.091	(0.288)	0.128	(0.335)
OTHERLOAN	0.095	(0.060)	0.101	(0.064)	0.096	(0.061)	0.097	(0.062)
RELPORT	37.81	(13.61)	38.64	(15.10)	37.74	(14.04)	42.15	(13.85)
ENFORCE	0.175	(0.023)	0.177	(0.026)	0.176	(0.024)	0.181	(0.026)
NEWBLOCK	670.3	(780.4)	769.5	(862.8)	690.3	(800.3)	823.6	(881.2)
GROWTH	0.047	(0.018)	0.049	(0.019)	0.047	(0.018)	0.050	(0.019)
D(1997)	0.350	(0.477)	0.398	(0.490)	0.362	(0.481)	0.401	(0.490)
D(1998)	0.151	(0.358)	0.241	(0.428)	0.167	(0.373)	0.311	(0.463)
Observations	14895		6372		19717		1550	

Table 5.12: Summary statistics by repayment behavior for first loans. Standard deviations are in parentheses.

	P(≥ 1 day late)			P(≥ 10 day late)		
	(1) 1st loans, 1st paym.	(2) 1st loans, final paym.	(3) All loans, all paym.	(4) 1st loans, final paym.	(5) All loans, all paym.	(6) All loans, , all paym. .
Personal Characteristics						
D(single)	0.071	0.052	0.040	0.078	0.042	0.061
	(2.32)*	(1.98)*	(4.38)**	(2.29)*	(2.99)**	(4.43)**
D(female)	-0.018	0.014	0.034	0.022	0.013	0.015
	(0.67)	(0.64)	(4.47)**	(0.75)	(1.10)	(1.25)
D(on black list)	0.485	0.296	0.277	0.329	0.376	0.381
	(3.33)**	(2.08)*	(6.82)**	(1.94)	(6.96)**	(7.06)**
ln(age)	-0.113	-0.145	-0.194	-0.126	-0.137	-0.172
	(2.46)*	(3.74)**	(14.57)**	(2.44)*	(6.48)**	(8.27)**
ln(non-bus. income)	8.58e-03	6.78e-03	6.96e-03	4.28e-03	1.10e-03	-8.17e-03
	(1.33)	(1.24)	(3.75)**	(0.59)	(0.38)	(2.87)**
ln(bus. income)	0.026	0.038	0.064	0.053	0.020	-0.016
	(1.08)	(1.81)	(8.00)**	(1.89)	(1.52)	(1.33)
Av. Arr. prior loan			0.101		0.053	0.051
			(13.55)**		(5.59)**	(5.47)**
(Av. Arr. pr. loan)^2			-2.47e-03		-8.43e-04	-7.92e-04
			(5.58)**		(2.38)*	(2.22)*
Max. Arr. prior loan			0.024		0.045	0.045
			(12.02)**		(10.27)**	(10.24)**
(Max. Arr. pr. loan)^2			-7.57e-04		-1.80e-03	-1.68e-03
			(10.55)**		(6.38)**	(5.94)**
(Max. Arr. pr. loan)^3			4.10e-06		2.44e-05	2.19e-05
			(7.95)**		(4.18)**	(3.71)**
(Max. Arr. pr. loan)^4			-4.44e-09		-1.05e-07	-9.27e-08
			(7.48)**		(3.07)**	(2.68)**
Arrears first paym.		0.144	0.156	0.134	0.140	0.142
		(14.46)**	(37.12)**	(13.93)**	(34.51)**	(35.08)**
(Arrears first paym.)^2		-2.15e-03	-2.74e-03	-1.81e-03	-2.22e-03	-2.26e-03
		(7.32)**	(14.30)**	(6.25)**	(13.08)**	(13.22)**
Arr. middle paym.		0.132	0.121	0.116	0.112	0.114
		(37.59)**	(38.29)**	(39.50)**	(41.85)**	(42.55)**
(Arr. middle paym.)^2		-8.84e-04	-9.40e-04	-7.91e-04	-8.30e-04	-8.36e-04
		(17.92)**	(11.79)**	(17.60)**	(11.68)**	(11.68)**
(Arr.1st)*(Arr.mid.)		-3.62e-03	-2.64e-03	-2.87e-03	-2.45e-03	-2.50e-03
		(7.01)**	(5.75)**	(6.72)**	(7.21)**	(7.45)**
Business Characteristics						
ln(assets)	-4.49e-03	0.012	7.59e-03	-0.053	-0.021	-0.042
	(0.23)	(0.69)	(1.31)	(2.33)*	(2.27)*	(4.61)**
Liabilities over assets	-0.139	-0.018	1.22e-03	-0.065	0.028	0.033
	(0.69)	(0.12)	(0.05)	(0.35)	(0.97)	(1.23)
D(other loans)	0.123	0.094	0.073	0.012	0.012	0.019
	(2.48)*	(2.24)*	(6.21)**	(0.21)	(0.66)	(1.03)
D(o. l.)*D(1999/2000)	0.126	0.070	0.039	0.179	0.123	0.112
	(1.64)	(1.18)	(1.85)	(2.36)*	(4.05)**	(3.73)**
ln(Business Age)	0.018	-0.010	0.014	8.82e-03	2.14e-03	7.03e-03
	(1.15)	(0.68)	(2.88)**	(0.46)	(0.26)	(0.87)
D(Commerce)	-0.032	-0.036	-0.041	-0.057	-0.028	-0.048
	(0.95)	(1.28)	(4.41)**	(1.49)	(1.97)*	(3.77)**
D(Service)	0.026	0.028	0.038	0.029	0.057	0.058
	(0.66)	(0.84)	(3.28)**	(0.66)	(3.20)**	(3.50)**

Table 5.13: Probit estimates for the probability of late payments.

Loan Characteristics						
ln(approved amount)	0.379	0.294	0.102	0.338	0.283	0.286
	(2.86)**	(2.75)**	(2.85)**	(2.26)*	(4.92)**	(5.46)**
ln(approved amount)^2	-0.038	-0.032	-0.014	-0.031	-0.025	-0.016
	(3.30)**	(3.45)**	(4.87)**	(2.40)*	(5.43)**	(3.92)**
Applied/appr. amount	0.041	0.027	0.032	0.021	0.047	
	(3.33)**	(2.54)*	(6.38)**	(1.41)	(6.93)**	
D(preferential loan)			-0.291		-0.644	
			(3.05)**		(2.92)**	
Interest Rate	-0.014	0.024	0.034	0.055	0.014	
	(0.31)	(0.63)	(2.72)**	(1.06)	(0.74)	
flat rate commission	0.017	-0.201	0.091	-0.254	-0.017	
	(0.16)	(2.13)*	(2.42)*	(2.12)*	(0.32)	
D(old penalty code)	-0.015	0.105	-0.010	2.97e-03	-0.025	
	(0.23)	(2.23)*	(0.62)	(0.05)	(1.03)	
D(weekly repayment)	-0.339	-0.322	-0.245	-0.495	-0.337	
	(2.81)**	(3.03)**	(7.82)**	(3.36)**	(6.69)**	
D(fortnightly repaym.)	-0.203	-0.105	-0.053	-0.135	-0.080	
	(2.95)**	(1.72)	(2.94)**	(1.64)	(2.83)**	
D(irregular repayment)	0.245	0.282	-0.173	-0.074	-0.103	
	(1.06)	(1.26)	(4.48)**	(0.22)	(1.73)	
Length of loan in days	-1.82e-04	4.40e-04	3.63e-04	3.03e-04	6.52e-04	
	(0.52)	(1.35)	(4.79)**	(0.72)	(5.86)**	
D(8-10 installments)	-5.14e-03	0.129	0.083	0.199	0.080	
	(0.11)	(3.29)**	(5.18)**	(3.61)**	(3.13)**	
D(11-14 installments)	0.014	0.231	0.137	0.308	0.120	
	(0.21)	(4.05)**	(7.19)**	(3.97)**	(4.00)**	
D(15-18 installments)	0.033	0.301	0.178	0.434	0.177	
	(0.35)	(3.69)**	(7.15)**	(3.90)**	(4.53)**	
D(19-24 installments)	0.049	0.385	0.204	0.491	0.210	
	(0.40)	(3.66)**	(6.85)**	(3.43)**	(4.46)**	
D(25-30 installments)	0.044	0.675	0.200	0.502	0.197	
	(0.22)	(4.18)**	(5.67)**	(2.29)*	(3.49)**	
D(>30 installments)	0.168	0.365	0.255	0.614	0.298	
	(0.69)	(1.72)	(5.82)**	(2.19)*	(4.28)**	
ln(val. chattel items)	0.014	-6.19e-03	0.003	7.83e-03	0.010	
	(1.99)*	(0.96)	(1.48)	(0.94)	(3.24)**	
D(Pers. Guarantee)	0.164	0.118	0.124	0.210	0.183	
	(6.38)**	(5.46)**	(17.36)**	(7.24)**	(16.29)**	
D(second loan)			0.025		-0.137	-0.127
			(2.52)*		(8.85)**	(8.47)**
D(third loan)			0.034		-0.200	-0.185
			(2.92)**		(10.70)**	(10.43)**
D(fourth loan)			0.018		-0.233	-0.215
			(1.36)		(10.64)**	(10.34)**
D(fifth loan)			4.91e-03		-0.263	-0.245
			(0.31)		(10.38)**	(10.11)**
D(sixth loan)			0.061		-0.208	-0.191
			(3.29)**		(7.07)**	(6.76)**
D(seventh loan)			0.058		-0.231	-0.219
			(2.63)**		(6.37)**	(6.23)**
D(eighth or higher l.)			0.047		-0.233	-0.239
			(2.33)*		(7.23)**	(7.73)**
D(middle payment)			0.507		0.617	0.592
			(59.17)**		(40.63)**	(39.42)**
D(final payment)			0.453		0.720	0.654
			(44.07)**		(42.47)**	(40.44)**

Table 13 continued: Probit estimates for the probability of late payments.

Environment						
D(Cochabamba)	0.068	0.137	0.013	0.318	0.089	0.049
	(0.81)	(1.97)*	(0.56)	(3.51)**	(2.44)*	(1.38)
D(Sucre)	-0.103	0.079	-0.048	-0.085	-0.222	-0.293
	(1.12)	(1.07)	(2.02)*	(0.86)	(5.92)**	(8.09)**
D(Trinidad)	-0.067	0.257	0.114	0.064	-0.039	-0.134
	(0.47)	(2.21)*	(2.86)**	(0.42)	(0.64)	(2.29)*
D(Tarija)	0.188	0.220	0.090	0.094	0.127	0.074
	(2.69)**	(3.84)**	(4.67)**	(1.27)	(4.52)**	(2.67)**
OTHERLOAN	0.030	-0.175	-0.378	-1.122	-0.494	-0.800
	(0.07)	(0.48)	(3.17)**	(2.32)*	(2.56)*	(4.24)**
RELPORT	-3.10e-03	2.69e-03	-2.83e-03	1.10e-03 -	-2.88e-03	-4.94e-03
	(1.17)	(1.28)	(4.05)**	(0.41)	(2.72)**	(4.93)**
ENFORCE	4.242	4.756	4.939	1.806	2.108	1.913
	(11.95)**	(18.07)**	(53.55)**	(5.42)**	(15.72)**	(14.78)**
NEWBLOCK	-6.30e-06	9.40e-06	3.80e-06	-1.34e-05	-1.39e-05	-8.70e-06
	(0.25)	(0.46)	(0.62)	(0.50)	(1.44)	(0.95)
GROWTH	-0.854	-0.280	-0.242	-0.500	-0.698	-0.336
	(0.89)	(0.33)	(0.96)	(0.44)	(1.74)	(0.88)
D(1997)	0.053	-0.117	-0.067	-0.079	-0.015	0.101
	(0.72)	(2.41)*	(4.27)**	(1.18)	(0.62)	(4.80)**
D(1998)	0.048	-0.118	-0.097	-0.022	0.006	0.226
	(0.49)	(1.63)	(4.19)**	(0.22)	(0.18)	(7.61)**
D(1999)	-0.069	-0.249	-0.161	-0.110	-0.027	0.277
	(0.52)	(2.74)**	(5.46)**	(0.91)	(0.60)	(7.37)**
D(2000)	-0.038	-0.236	-0.071	7.91e-03	0.048	0.449
	(0.23)	(1.90)	(1.82)	(0.05)	(0.78)	(8.54)**
D(2nd quarter)	-0.011	0.030	0.016	-0.072	-0.049	-0.021
	(0.32)	(0.98)	(1.66)	(1.84)	(3.13)**	(1.41)
D(3rd quarter)	-0.048	-0.146	-0.075	-0.206	-0.090	-0.041
	(1.14)	(4.13)**	(6.72)**	(4.42)**	(5.10)**	(2.39)*
D(4th quarter)	-0.107	-0.116	-0.065	-0.108	-0.083	-0.015
	(2.74)**	(2.91)**	(5.51)**	(2.06)*	(4.38)**	(0.87)
Constant	-2.525	-2.026	-2.083	-2.378	-3.032	-2.431
	(5.30)**	(5.09)**	(14.94)**	(4.36)**	(13.64)**	(13.20)**
Observations	21680	20039	186729	20039	186729	186729
Percentage late/default	10.1	31.32	26.0	11.18	6.2	6.2
Pseudo-R^2	0.04	0.16	0.14	0.24	0.23	0.23
Chi2 (LR/Wald)	588.73	4061.13	19180.89	3357.57	13304.96	12781.50
Degrees of Freedom	47	52	67	52	68	51
Log Likelihood	-6797.28	-10500.57	-91843.50	-5393.09	-33153.08	-33502.95

Table 13 continued: Probit estimates for the probability of late payments for first and final payments, column (6) excludes potentially endogenous variables. Robust standard errors are in parentheses, ** denotes significant at the 1% level, * at the 5% level. Sample: A 70% random sample of all loans with $\geq$ 6 installments, all installments due between January 1996 and June 2000.

	first payment	middle payment	final payment
First loans			
$\geq$ 1 day late	10.03	25.29	31.32
$\geq$ 10 days late	1.10	6.53	11.18
All loans			
$\geq$ 1 day late	13.91	28.81	32.81
$\geq$ 10 days late	1.12	6.25	11.32

Table 5.14: Percentage of late payments for first, middle, and final payments for first loans of new clients and for all loans together. Calculations are based on the sample used for the regression analyses, that is, loans with at least six scheduled repayments, payments in 1996 to 2000, and corresponding balance observations in the non-agricultural sectors.

	p10		**p20**		**p30**		
late1	0	1	0	1	0	1	**Total**
0	7,943	7,433	12,278	3,098	13,580	1,796	15,376
	(42.9)	(40.1)	(66.3)	(16.7)	(73.3	(9.7	(83.0)
1	488	2,664	1,149	2,003	1,572	1,580	3,152
	(2.6)	(14.4)	(6.2)	(10.8)	(8.5	(8.5	(17.0)
Total	8,431	10,097	13,427	5,101	15,152	3,376	18,528
	(45.5)	(54.5)	(72.5)	(27.5)	(81.8	(18.2	(100.0)

Table 5.15: Out-of-sample predictive power for late payments. Calculations are based on a pooled probit estimate for the probability that a payment (first, middle or final) is at least one day late using data from 1996 to 1999, predictions are made for 2000. Percentages are in parentheses.

	First loans. first paym.	All loans. all paym.
Personal Characteristics		
D(single)	-1.786	-1.557
	(0.99)	(3.64)**
D(female)	0.452	-0.581
	(0.30)	(1.72)
D(on black list)	-9.409	-14.723
	(0.95)	(6.26)**
ln(age)	2.205	5.951
	(0.85)	(10.26)**
ln(non-bus. income)	-0.127	-0.129
	(0.34)	(1.54)
ln(bus. income)	0.180	-1.039
	(0.13)	(3.06)**
Av. Arr. prior loan		-2.925
		(8.92)**
(Av. Arr. pr. loan)^2		0.039
		(4.50)**
Max. Arr. prior loan		-0.996
		(7.34)**
(Max. Arr. pr. loan)^2		0.020
		(3.85)**
(Max. Arr. pr. loan)^3		-9.21e-04
		(2.45)*
(Max. Arr. pr. loan)^4		9.79e-07
		(1.88)
Arrears first paym.		-6.027
		(18.72)**
(Arrears first paym.)^2		0.077
		(7.60)**
Arr. middle paym.		-5.935
		(38.63)**
(Arr. middle paym.)^2		0.036
		(14.34)**
(Arr.1st)*(Arr.mid.)		0.099
		(8.54)**
Business Characteristics		
ln(assets)	-0.375	0.136
	(0.34)	(0.57)
Liabilities over assets	3.154	0.010
	(0.34)	(0.01)
D(other loans)	-2.191	-1.321
	(0.79)	(2.37)*
D(o. l.)*D(1999/2000)	-0.935	-2.248
	(0.23)	(2.74)**
ln(Business Age)	0.079	-0.224
	(0.09)	(1.02)
D(Commerce)	-0.623	1.224
	(0.33)	(2.96)**
D(Service)	-1.404	-2.321
	(0.62)	(4.39)**

	First loans. first paym.	All loans. all paym.
Loan Characteristics		
ln(approved amount)	-6.886	-7.629
	(0.94)	(5.06)**
ln(approved amount)^2	0.746	0.732
	(1.18)	(5.91)**
Applied/appr. amount	-1.011	-1.780
	(1.30)	(7.12)**
D(preferential loan)	-1.802	1.529
	(0.03)	(1.20)
Interest Rate	1.447	-0.155
	(0.54)	(0.28)
flat rate commission	0.628	-0.695
	(0.10)	(0.45)
D(old penalty code)	2.018	-0.562
	(0.50)	(0.72)
D(weekly repayment)	15.100	12.170
	(2.24)*	(9.47)**
D(fortnightly repaym.)	10.158	5.205
	(2.61)**	(6.95)**
D(irregular repaym.)	-4.673	5.749
	(0.33)	(4.20)**
Length of loan in days	0.017	-0.017
	(0.89)	(5.64)**
D(8-10 installments)	-0.414	-3.054
	(0.16)	(4.51)**
D(11-14 installments)	-1.632	-4.213
	(0.45)	(5.27)**
D(15-18 installments)	-2.915	-6.311
	(0.55)	(6.03)**
D(19-24 installments)	-3.424	-7.434
	(0.50)	(6.03)**
D(25-30 installments)	-5.075	-7.366
	(0.47)	(4.95)**
D(>30 installments)	-8.761	-9.386
	(0.64)	(5.19)**
ln(val. chattel items)	-0.409	-0.289
	(0.97)	(3.26)**
D(Pers. Guarantee)	-3.579	-6.406
	(2.42)*	(20.01)**
D(second loan)		1.747
		(4.08)**
D(third loan)		2.912
		(5.70)**
D(fourth loan)		3.676
		(6.34)**
D(fifth loan)		4.357
		(6.34)**
D(sixth loan)		3.314
		(4.10)**
D(seventh loan)		1.932
		(1.93)
D(eighth or higher l.)		2.805
		(3.23)**

Table 5.16: Cox proportional hazard estimates for the duration of arrears.

Environment		
D(Cochabamba)	-3.576	-2.994
	(0.72)	(2.99)**
D(Sucre)	3.693	6.503
	(0.68)	(6.36)**
D(Trinidad)	6.536	1.542
	(0.78)	(0.91)
D(Tarija)	-4.714	-3.521
	(1.10)	(3.77)**
OTHERLOAN	5.370	13.745
	(0.21)	(2.79)**
RELPORT	0.074	0.098
	(0.47)	(3.17)**
ENFORCE	-52.601	-91.874
	(2.70)**	(22.47)**
NEWBLOCK	-4.20e-04	3.78e-02
	(0.03)	(1.30)
GROWTH	23.711	11.637
	(0.41)	(0.98)
D(2nd quarter)	0.437	1.385
	(0.22)	(3.26)**
D(3rd quarter)	1.668	2.668
	(0.68)	(5.54)**
D(4th quarter)	2.262	2.758
	(1.00)	(5.50)**
Observations	21683	186729
Chi2 (LR/Wald)	58.80	6386.57
Degrees of Freedom	44	61
Log Likelihood	-176279.33	-1579305.22

Table 16 continued: Cox proportional hazard estimates for the duration of arrears for first and final payments. The Cox estimates are listed in percentage terms (multiplied by 100). Stratification: by year for the first payment. Joint estimates (last column): additional strata for being late in the first and/or middle payment and for middle and final payments. Robust standard errors are in parentheses, ** denotes significant at the 1% level, * at the 5% level. Sample: A random sample of 70% of all loans with ≥ 6 installments, all installments due between January 1996 and June 2000.

	All loans, all payments			
	punctual payments		late payments	
Personal Characteristics				
D(single)	0.179	*(0.383)*	0.199	*(0.400)*
D(female)	0.638	*(0.481)*	0.643	*(0.479)*
D(on black list)	5.69e-03	*(0.075)*	9.17e-03	*(0.095)*
ln(age)	3.634	*(0.281)*	3.610	*(0.271)*
ln(non-business income)	2.611	*(2.283)*	2.538	*(2.302)*
ln(business income)	5.070	*(0.792)*	5.136	*(0.777)*
Av. Arrears previous loan	0.329	*(1.377)*	0.568	*(2.169)*
Max Arrears previous loan	1.684	*(5.190)*	2.588	*(7.263)*
Days in Arrears first paym.	0.149	*(1.033)*	0.912	*(3.220)*
Days in Arr. middle paym.	0.189	*(1.511)*	1.841	*(6.040)*
Business Characteristics				
ln(assets)	7.234	*(1.175)*	7.274	*(1.155)*
Liabilities over assets	0.033	*(0.154)*	0.030	*(0.090)*
D(other loan)	0.177	*(0.382)*	0.176	*(0.381)*
D(other loan)*D(1999/2000)	0.081	*(0.273)*	0.052	*(0.223)*
ln(Business Age)	1.763	*(0.803)*	1.802	*(0.796)*
D(Commerce)	0.579	*(0.494)*	0.556	*(0.497)*
D(Service)	0.150	*(0.357)*	0.170	*(0.375)*
Observations	143390		49132	

Table 5.17: Summary statistics for punctual and late payments for first, middle, and final payments of all loans with $\geq$ six installments.

	All loans, all payments			
	punctual payments		late payments	
Loan Characteristics				
ln(approved amount)	5.851	*(0.966)*	5.872	*(0.947)*
Applied over appr. amount	1.253	*(0.682)*	1.298	*(0.665)*
D(preferential loan)	0.016	*(0.124)*	1.70e-03	*(0.041)*
Interest Rate	3.192	*(0.439)*	3.190	*(0.409)*
flat rate commission	0.998	*(0.095)*	0.998	*(0.106)*
D(old penalty code)	0.318	*(0.466)*	0.340	*(0.474)*
D(weekly)	0.124	*(0.330)*	0.105	*(0.306)*
D(fortnightly)	0.353	*(0.478)*	0.405	*(0.491)*
D(irregular)	0.041	*(0.199)*	0.013	*(0.115)*
Length of loan in days	299.6	*(162.8)*	288.9	*(138.0)*
D(8-10 installments)	0.212	*(0.409)*	0.216	*(0.411)*
D(11-14 installments)	0.326	*(0.469)*	0.328	*(0.470)*
D(15-18 installments)	0.161	*(0.368)*	0.161	*(0.368)*
D(19-24 installments)	0.130	*(0.336)*	0.140	*(0.347)*
D(25-30 installments)	0.056	*(0.230)*	0.053	*(0.224)*
D($>$30 installments)	0.042	*(0.202)*	0.036	*(0.187)*
ln(value of chattel items)	5.799	*(2.148)*	5.884	*(2.137)*
D(pers. Guarantee)	0.463	*(0.499)*	0.510	*(0.500)*
Environment				
Cochabamba	0.105	*(0.307)*	0.124	*(0.330)*
Sucre	0.111	*(0.314)*	0.147	*(0.354)*
Trinidad	0.034	*(0.182)*	0.043	*(0.203)*
Tarija	0.072	*(0.259)*	0.076	*(0.265)*
OTHERLOAN	0.126	*(0.078)*	0.109	*(0.071)*
RELPORT	48.52	*(20.80)*	42.42	*(16.42)*
ENFORCE	0.151	*(0.069)*	0.179	*(0.048)*
NEWBLOCK	969.9	*(804.3)*	853.0	*(817.0)*
GROWTH	0.039	*(0.023)*	0.044	*(0.022)*
D(1997)	0.240	*(0.427)*	0.289	*(0.453)*
D(1998)	0.196	*(0.397)*	0.285	*(0.451)*
D(1999)	0.171	*(0.376)*	0.142	*(0.349)*
D(2000)	0.158	*(0.365)*	0.077	*(0.267)*
D(2nd Quarter)	0.274	*(0.446)*	0.285	*(0.451)*
D(3rd Quarter)	0.270	*(0.444)*	0.228	*(0.419)*
D(4th Quarter)	0.225	*(0.418)*	0.232	*(0.422)*
Observations	143390		49132	

Table 17 continued:Summary statistics for punctual and late payments for first, middle, and final payments of all loans with $\geq$ six installments.

Bibliography

AIYAGARI, S. R. (1995): "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 103(6), 1158–75.

ALEEM, I. (1993): "Imperfect Information, Screening, and the Costs of Informal Lending: A Study of a Rural Credit Market in Pakistan," in *The Economics of Rural Organization: Theory, Practice, and Policy*, ed. by K. Hoff, A. Braverman, and J. E. Stiglitz, chap. 7, pp. 87–108. Oxford University Press, New York.

ALLEN, F. (1981): "The Prevention of Default," *Journal of Finance*, 36(2), 271–76.

ARMENDÁRIZ DE AGHION, B., AND J. MORDUCH (2000): "Microfinance Beyond Group Lending," *Economics of Transition*, 8(2), 401–20.

ASOFIN (2000): "Boletin Financiero, Microfinanzas," various issues, La Paz, Bolivia.

ATKINSON, A. B. (1977): "Optimal Taxation and the Direct versus Indirect Tax Controversy," *Canadian Journal of Economics*, 10, 590–606.

ATKINSON, A. B., AND J. STIGLITZ (1976): "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics*, 6, 55–75.

BASSETTO, M. (1999): "Optimal Fiscal Policy with Heterogeneous Agents," unpublished manuscript, University of Minnesota.

BESLEY, T., AND A. R. LEVENSON (1996): "The Role of Informal Finance in Household Capital Accumulation: Evidence from Taiwan," *The Economic Journal*, 106, 39–59.

Bolton, P., and X. Freixas (2000): "Equity, Bonds and Bank Debt: Capital Structure and Financial Market Equilibrium under Asymmetric Information," *Journal of Political Economy*, 108(2), 324–51.

Bolton, P., and D. S. Sharfstein (1990): "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, 80, 93–106.

Boyes, W. J., D. L. Hoffman, and S. A. Low (1989): "An Econometric Analysis of the Bank Credit Scoring Problem," *Journal of Econometrics*, 40(1), 3–15.

Bulow, J., and K. Rogoff (1989): "Sovereign Debt: Is to Forgive to Forget?," *American Economic Review*, 79(1), 43–50.

Caja Los Andes (1999): *Memoria Anual.* La Paz, Bolivia.

Chamley, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54(3), 607–22.

——— (2001): "Capital Income Taxation, Wealth Distribution and Borrowing Constraints," *Journal of Public Economics*, 79, 55–69.

Chari, V. V., and P. J. Kehoe (1999): "Optimal Fiscal and Monetary Policy," *NBER Working Paper*, #6891.

Chaudhury, I. A., and I. Matin (2001): "Dimensions and Dynamics of Microfinance Membership Overlap: A Micro Study from Bangladesh," unpublished manuscript, Research and Evaluation Division, BRAC, Bangladesh.

Churchill, C. F. (1999): *Client-Focused Lending: The Art of Individual Lending.* Calmeadow, Toronto.

CIA (1999): "CIA Factbook: Bolivia," Washington, D.C.

Coleman, B. E. (1999): "The Impact of Group Lending in Northeast Thailand," *Journal of Development Economics*, 60, 105–141.

——— (2001): "Microfinance in Northeast Thailand: Who Benefits and How Much?," unpublished manuscript, Asian Development Bank.

Conning, J. (2000): "Joint Liability, Peer Monitoring, and the Making of Social Collateral," unpublished manuscript, Williams College and Yale University.

Copestake, J., S. Bhalotra, and S. Johnson (2001): "Assessing the Impact of Microcredit: A Zambian Case Study," *Journal of Development Studies*, 37(4), 81–100.

Cremer, H., P. Pestieau, and J. C. Rochet (1999): "Direct versus Indirect Taxation: The Design of the Tax Structure Revisited," Discussion Paper 9910, CORE, Universite Catholique de Louvain.

De Soto, H. (1989): *The Other Path.* I.B.Tauris & Co Ltd, London.

Deaton, A. (1979): "Optimally Uniform Commodity Taxes," *Economic Letters*, 2, 357–61.

Deininger, K., and L. Squire (1996): "Measuring Income Inequality: A New Database," Worldbank.

Domeij, D., and J. Heathcote (2000): "Capital versus Labor Income Taxation with Heterogeneous Agents," unpublished manuscript, Stockholm School of Economics.

Eaton, J., M. Gersovitz, and J. E. Stiglitz (1986): "The Pure Theory of Country Risk," *European Economic Review*, 30, 481–513.

Evans, D. S. (1987a): "The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries," *Journal of Industrial Economics*, 35 (4), 567–81.

——— (1987b): "Tests of Alternative Theories of Firm Growth," *Journal of Political Economy*, 95(4), 657–74.

Frankfurter Rundschau (1998): "Auch Kleinfirmen machen Mist," Nr. 241, October 17th, p. 15.

Ghosh, P., and D. Ray (1999): "Information and Enforcement in Informal Credit Markets," Discussion Paper 93, Institute for Economic Development, Boston University.

Gonzalez-Vega, C., R. L. Meyer, S. Navajas, M. Schreiner, J. Rodriguez-Meza, and G. F. Monje (1996): "Microfinance Market Niches and Client Profiles in Bolivia," Ohio State University, Rural Finance Program, Economics and Sociology Occasional Paper #2346.

Greene, W. H. (1998): "Sample Selection in Credit-Scoring Models," *Japan and the World Economy*, 10, 299–316.

——— (2000): *Econometric Analysis, 4th Edition.* Prentice Hall.

Gromb, D. (2001): "Dynamic Financial Contracting and Renegotiation," unpublished manuscript, MIT and CEPR.

Heckman, J. (1976): "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement*, 5, 479–92.

Hellwig, M. F. (1977): "A Model of Borrowing and Lending with Bankruptcy," *Econometrica*, 45 (8), 1897–906.

Hoff, K., A. Braverman, and J. E. Stiglitz (1993): *The Economics of Rural Organization: Theory, Practice, and Policy.* Oxford University Press, New York.

Holmström, B., and J. Tirole (1997): "Financial Intermediation, Loanable Funds, and the Real Sector.," *Quarterly Journal of Economics*, 112(3), 663–91.

Inter-American Development Bank (1998): "Microenterprise Development Review 1(1)," Sustainable Development Department, Microenterprise Unit.

——— (1999): "Bolivia: Basic Socio-Economic Data," Statistics and Quantitative Analysis Unit.

IPC GmbH (2000): "Balance Information and Efficiency Indicators," *http://www.ipcgmbh.de*.

Jones, L. E., R. E. Manuelli, and P. E. Rossi (1997): "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 73, 93–117.

Judd, K. L. (1985): "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28, 59–83.

——— (1997): "The Optimal Tax Rate for Capital Income is Negative," *NBER Working Paper*, #6004.

Kan, K. (2000): "Informal Capital Sources and Household Investment: Evidence from Taiwan," *Journal of Development Economics*, 62, 209–32.

Kelly, M. (2000): "Inequality and Crime," *Review of Economics and Statistics*, 82(4), 530–39.

Kervyn, B. (2001): "Evaluation and Progress Report of the 2 Community Development Projects," Viet Nam Plus, 19/30 Tran Binh Trong, P5 Binh Thanh, HCMC, Vietnam.

Khandker, S. R., H. A. Samad, and Z. H. Khan (1998): "Income and Employment Effects of Micro-Credit Programmes: Village-Level Evidence from Bangladesh," *Journal of Development Studies*, 35(2), 96–124.

Kiefer, N. M. (1998): "Economic Duration Data and Hazard Functions," *Journal of Economic Literature*, 26, 646–79.

Klein, P., and J.-V. Rios-Rull (1999): "Time Consistent Optimal Fiscal Policy," unpublished manuscript, Institute for International Economic Studies, Stockholm University.

KOCHAR, A. (1997): "An Empirical Investigation of Rationing Constraints in Rural Credit Markets in India," *Journal of Development Economics*, 53, 339–71.

KRAHNEN, J. P., AND R. H. SCHMIDT (1994): *Development Finance as Institution Building.* Westview Press, Boulder, Colorado.

MADAJEWICZ, M. (1999): "Does the Credit Contract Matter? The Impact of Lending Programs on Poverty in Bangladesh," unpublished manuscript, Columbia University.

MADDALA, G. S. (1983): *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge University Press.

MATIN, I. (2001): "Dimensions and Dynamics of MFI Competition in Bangladesh," CGAP Note Book 1, Consultative Group to Assist the Poorest.

MCKERNAN, S.-M. (2000): "The Impact of Micro-Credit Programs on Self-Employment Profits: Do Non-Credit Program Aspects Matter?," unpublished manuscript, The Urban Institute.

MCFADDEN, D. (1976): "A Comment on Discriminant Analysis "versus" Logit Analysis," *Annals of Economic and Social Measurement*, 5(4), 511–23.

MCPHERSON, M. A. (1996): "Growth of Micro and Small Enterprises in Southern Africa," *Journal of Development Economics*, 48, 253–77.

MONTIEL, P. J., P. R. AGENOR, AND N. UL HAQUE (1993): *Informal Financial Markets in Developing Countries.* Blackwell, Oxford.

MORDUCH, J. (1998): "Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh," unpublished manuscript, Princeton University.

——— (2000): "The Microfinance Promise," *Journal of Economic Literature*, 37, 1569–614.

MOSLEY, P. (2001): "Microfinance and Poverty in Bolivia," *Journal of Development Studies*, 37(4), 101–32.

MURINDE, V. (1996): *Development Banking and Finance*. Avebury, Aldershot.

NAITO, H. (1999): "Re-Examination of Uniform Commodity Taxes under a Non-Linear Income Tax System and its Implication for Production Efficiency," *Journal of Public Economics*, 71, 165–88.

NAVAJAS, S., J. CONNING, AND C. GONZALEZ-VEGA (1999): "Lending Technologies, Competition, and Consolidation in the Market for Microfinance in Bolivia," unpublished manuscript, Rural Finance Group, Ohio State University.

NAVAJAS, S., M. SCHREINER, R. L. MEYER, C. GONZALEZ-VEGA, AND J. ROGRIGUEZ-MEZA (2000): "Microcredit and the Poorest of the Poor: Theory and Evidence from Bolivia," *World Development*, 28(2), 333–46.

ORLANDO, M. B., AND M. POLLACK (2000): "Microenterprises and Poverty: Evidence from Latin America," Discussion paper, Inter-American Development Bank, Sustainable Development Department, Microenterprise Unit.

PITT, M. M. (1999): "The Effect of Non-Agricultural Self-Employment Credit on Contractual Relations and Employment in Agriculture: The Case of Microcredit Programs in Bangladesh," unpublished manuscript, Brown University, forthcoming in Bangladesh Development Studies.

PITT, M. M., AND S. R. KHANDKER (1998): "The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?," *Journal of Political Economy*, 106(5), 958–96.

PITT, M. M., S. R. KHANDKER, O. H. CHOWDHURY, AND D. MILLIMET (1998): "Credit Programs for the Poor and the Nutritional Status of Children in Rural Bangladesh," Discussion Paper 98/04, Brown University, Department of Economics.

PRESS, S. J., AND S. WILSON (1978): "Choosing Between Logistic Regression and Discriminant Analysis," *Journal of the American Statistical Association*, 73, 699–705.

RAMSEY, F. (1927): "A Contribution to the Theory of Taxation," *Economic Journal*, 37, 47–61.

RHYNE, E. (2001): *Mainstreaming Microfinance: How Lending to the Poor Began, Grew, and Came of Age in Bolivia*. Kumarian Press, Bloomfield, Connecticut.

SCHREINER, M. (1999): "A Scoring Model of the Risk of Costly Arrears at a Microfinance Lender in Bolivia," unpublished manuscript, Center for Social Development, Washington University in St. Louis.

SHESHINSKI, E. (1972): "The Optimal Linear Income-Tax," *Review of Economic Studies*, 39, 297–302.

STIGLITZ, J. E. (1985): "Inequality and Capital Taxation," Discussion Paper 457, IMSSS, Stanford University.

THE ECONOMIST (2001a): "Does Inequality Matter?," June 16th.

——— (2001b): "The Word Bank's Report on Institutions," September 15th.

THE MICROBANKING BULLETIN (2000): Issue No. 5, Calmeadow, Toronto.

TUNALI, I. (1983): "A Common Structure for Models of Double Selection," Discussion Paper 8304, Social Systems Research Institute, University of Wisconsin-Madison.

VELLA, F. (1998): "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, 33, 163–83.

VILLAS-BOAS, J. M., AND U. SCHMIDT-MOHR (1999): "Oligopoly with Asymmetric Information: Differentiation in Credit Markets," *Rand Journal of Economics*, 30(3), 375–96.

Von Stauffenberg, D. (2001): "How Microfinance Evolves: What Bolivia Can Teach Us," *Microenterprise Development Review*, 4(1).

Worldbank (2000a): "Bolivia at a Glance," Washington D.C.

——— (2000b): "Development Goals: Poverty," Washington D. C.

Hiermit erkläre ich, daß ich die Dissertation selbständig angefertigt habe und mich anderer als der in ihr angegebenen Hilfsmittel nicht bedient habe, insbesondere, daß aus anderen Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche gekennzeichnet sind und mit Quellenangaben versehen sind, nicht stattgefunden haben.

Stuttgart, im April 2002