

D
Dipartimento

S
Scienze

E
Economiche

Working Paper

Department
of Economics

Ca' Foscari University of
Venice

Charles Grant
Mario Padula

Bounds on repayment
behavior: evidence for the
consumer credit market



Bounds on repayment behavior: evidence for the consumer credit market

Charles Grant

University of Reading

Mario Padula

Ca' Foscari University of Venice and CSEF

First Draft: November 2007

Abstract

How does the punishment for default affect repayment behavior? We use administrative data, provided by the leading Italian lender of unsecured credit to the household sector, to analyze households repayment behavior. Administrative data are particularly well suited to study what factors are responsible for default, but raise a fundamental econometric problem, since they identify the determinants of repayment behavior only for those who are granted credit. To overcome this problem, we provide upper and lower bounds on the determinants of repayment behavior. Moreover, we show how to use the restrictions from the theory to narrow the bounds.

Mario Padula gratefully acknowledges financial support of the Finance and Consumption Chair in the European Community which enabled this project to be completed. We would like to thank for comments and suggestions at various stages of this project Alberto Bennardo, Giuseppe Bertola, Agar Brugiavini, Richard Disney, John Duca, Burcu Duygan, Luigi Guiso, Michael Haliassos, Stefan Hochguertel, Winfried Koeninger, Tullio Jappelli, Theresa Osborne, Frank Vella and Guglielmo Weber, and many other seminar participants. All remaining errors, are of course our own.

Keywords

Manski Bounds, Consumer Credit, Default

JEL Codes

D14, K42, O17

Address for correspondence:

Mario Padula

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy

Phone: (+39) 041 2349184

Fax: (+39) 041 2349176

e-mail: mpadula@unive.it

This Working Paper is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.

1 Introduction

Which households repay their debts and which borrowers default? Which factors affect repayment behavior, and how large are the effects? This paper addresses these questions and provides some evidence using data provided by the leading lender of unsecured credit to the Italian household sector. Credit markets in Italy are small by EU standards, but they grew quite rapidly over the past 20 years, as well documented by Casolaro, Gambacorta and Guiso (2005). The trend is similar for consumer credit (e.g. non-housing debt), which accounts for 8.1 percent of the GDP in 2003 and is largely unsecured. Theory predicts that incentives to repay depend crucially on how default is punished.

This paper provides a methodology to deal with the econometric issues that arise when one uses lenders' data to investigate the factor affecting repayment behavior. Lenders' data have a number of advantages. Since the data record the repayment history of applicants who were given credit, this allows us to observe default, which is a rare event in general household surveys. Even on the few occasions appropriate questions on default are included, this is likely to be underreported.¹ A second advantage is that administrative data records all the variables that affect the decision to lend, while survey data typically has only a subset of them. However, lenders' data typically allow to identify only the repayment behavior for those who are granted credit. To the extent that non-rejected applicants are a selected sample from the population of applicants, one cannot draw reliable policy implications from the observation of non-

¹For example, Fay, Hurst and White (2002) found that only around 250 US households reported filing for bankruptcy in the 1996 wave of the PSID, around half the national filing rate. Moreover, and more seriously, only a small proportion of households in serious arrears subsequently files for bankruptcy.

rejected applicants only.

The standard way of coping with the issue is to find some exclusion restrictions that affect the rejection rule but not the repayment behavior. However, these exclusion restrictions are often implausible for lenders' data, since the variables recorded in the administrative archives are used to screen the applicants on the grounds that they predict the chances of default and therefore repayment behavior. To overcome such difficulty, we describe an alternative identification method which places upper and lower bounds on the true effect. We also show how simple, and plausible, assumptions about the behaviour of the lender (that they lend to good risk and refuse credit to bad risks) can narrow the bounds on the estimates.

This study uses a novel data-set drawn directly from the administrative records of the leading lender of unsecured credit to the Italian household sector. The data provide detailed information on the characteristics of contracts, customers, repayment and, importantly, rejected applications. This information is crucial if one wants to draw inference on household repayment behavior, and to account for the fact that households granted credit are likely to be different from and those which are refused.

The paper analyzes the role of several factors affecting repayment behavior and also focus on how the quality of judicial enforcement and the availability of informal credit markets affect households' default on consumer credit contracts.² The incentives for individuals to default depends on the penalty incurred when not repaying. If the debt is collateralized,

²An issue left for future research is to study the effect of information sharing through credit bureaus on repayment. To investigate such an issue one needs using data coming from lenders adopting different information sharing policies.

and the debt is not repayed, then the property pledged as collateral is transferred to the creditor. The speed with which the asset is transferred depends, among others things, on how long it takes for the court to enforce the contract. Thus, a more tardy enforcement of debts enforcing makes it less costly for the borrower to fail to repay. But a slow and costly judicial enforcement procedure can also have an important role when the debt is not collateralized, in that it might discourage lenders from court action should the borrower default.

Using data on civil trials provided by the Italian Institute of Statistics (ISTAT), Jappelli, Pagano and Bianco (2004) and Fabbri and Padula (2005) document the effect of judicial enforcement on credit to firms and to households, respectively. Both studies exploit the large variation across Italy in judicial enforcement, and find significant effects on access to credit. Our approach is similar, with variation in the quality of judicial institutions playing a major role in the analysis. However, in contrast to these papers, the focus is on consumer credit market, which is mostly unsecured. Moreover, we investigate the effect of judicial enforcement and of informal credit markets on borrowers' repayment behavior and not on borrowing restrictions.

The value of non-housing consumer loans is typically small and poorly collateralized (even for installment credit, the resale value of the good would rarely cover the outstanding debt). Failure to pay, in such circumstances, often only amounts to being blacklisted and excluded from future borrowing in formal credit markets.³ The cost of exclusion will thus depend on whether alternative sources of credit are available. Households with access to alternative sources of credit are likely to view the deter-

³Unlike the US, institutions that specialize in subprime lending to households with poor credit histories are absent in Italy.

rent effect of exclusion from the formal sector less seriously, and thus have lower incentives to repay debts incurred there. Among alternative credit providers, an important role is played by informal credit markets: around 3% of Italian households are indebted to family and friends. This is not a negligible number: just below 11% of Italian households are indebted with banks and other financial institutions.⁴

Other than affecting the borrowers' outside option, there is another important way in which financial help from relatives or friends might affect repayment behavior. As emphasized in a number of papers (see for instance Ghosh, Mookherjee and Ray, 2000), family networks have access to a better monitoring technology, and thus can lend to borrowers who would otherwise be unable to pay their debts in the formal sector. This can enable financially troubled borrowers to meet their payments on consumer credit contracts even when their assets would not otherwise allow them to do so, and consequently makes default less likely. The paper will thus investigate whether the availability of family and friends' financial help makes borrower default more, or less, likely.

The rest of the paper is organized as follows. Section 2 reviews some background literature. The data are described in Section 3, while Section 4 deals with the selection issues involved in the use of lenders' data. Section 5 discusses parametric and semi-parametric estimation. Section 6 presents the results and 7 concludes.

⁴Moreover, data from the Consumer Credit and Lending to Households in Europe - ECRI 2005 Statistical Package reveal that between 2000 and 2004 the ratio between consumer credit and disposable income is just above 5%; in the Bank of Italy Survey of Household Income and Wealth such ratio is around 4%.

2 Background literature

The early literature on the lender-borrower relationship showed how asymmetric information could cause banks and other lenders to restrict access to credit (see, for instance, the pioneering work of Stiglitz and Weiss, 1981). In these models, default is exogenous but creditors can not tell *a priori* which agents will default and which will not, hence they offer the same contract to all borrowers. In contrast, Jaffee and Russell (1976) discussed how agents who do not bear the full consequences of their actions may indulge in riskier behaviour and thus are more likely to default on their debts. While this literature considered entrepreneurs, the insights are also relevant for consumption smoothing. In a more recent literature, Kehoe and Levine (1993) and Kocherlakota (1996) explicitly model the decision to default of infinitely lived consumers in a general equilibrium framework. Whether households default depends on how severely they are punished: agents compare the punishment for default with the gain from not repaying their debts. In these models, default is punished by autarky; permanent exclusion from borrowing and saving in all future time periods. These studies show that credit constraints would arise endogenously since above some maximum level of debt, default is assured, and it is never rational for lenders to extend credit beyond this level. A key point in this literature is that incentives to default depend on the agents' outside option. In empirical studies this is important since debtors can not be permanently excluded from credit markets. In the US, for example, bankruptcy can not be recorded in credit files for more than 10 years, and in practice, these households gain access to credit and to saving instruments much more quickly. Moreover, informal credit channels, such as friends and family, may be available to such households

without interruption. Incentives to default will also depend on the cost of enforcing debts, in terms of both money and the time it takes for lenders to recover their debts. Making the enforcement of debt contracts more difficult makes lending less attractive for lenders. This paper will explicitly deal with these two issues.

While the relation between credit contracts and both the legal rules and their enforcement is widely studied (see for instance Fay, Hurst and White, 2002, Grant, 2003, and Fabbri and Padula, 2004), much less it has been said on the effect of non-market sources of credit. In a very different context, Banerjee and Newman (1998) show that these alternative credit sources can have important effects on development. They argue that in the formal sector, informational asymmetries can be large, while they are much smaller in the informal sector in which agents behaviour can be much more easily observed: friends and family are likely to know whether people they know closely are reliable and will repay their debts. In contrast, our paper provides empirical evidence (albeit on credit to consumers not producers) showing that those types of household for which informal credit is more common are less likely to repay their debts in the formal sector, everything else being equal.

3 Data

We take data from three different sources, described in turn. For information on borrowing, we have a unique data set which consists of a random sample of households that are in the full administrative database of Findomestic Banca for 1996 - 1999. Findomestic specializes in non-mortgage lending to the Italian household sector, much of it (61 percent in our sample) via instalment credit made available by the retailer at the

point of sale. The bank also supplies revolving credit (37 percent of the contracts in our sample) in the form of credit cards. Lastly, and rather less importantly, the bank offers personal loans, a market which it has entered more recently. Our lender is the market leader for these types of credit in Italy. In the data, the median debt is only 700 euros. Even though much of the debt is instalment credit, in practice the recovery value of the good which was purchased is small, hence our lender treats all the loans as unsecured, and does not attempt to repossess the good.⁵

In 1999, the last year for which we have data, our lender had three million customers in their credit records. From this they have provided a random sample of approximately 120,000 clients, and since clients may have more than one contract, we have information on roughly 200,000 contracts. The lender has made all its customer information available to us, except the specific credit score for the customer. The lender's aim in collecting the data is to identify suitable consumers with which to build long-term relationships, such as middle-income families with steady jobs. The data records all applications that have been made by the household; whether credit was granted; and the repayment history of the household for each debt contract. The data also includes information about the household's characteristics, such as date of birth of the head and of the spouse, the profession of household members, the province

⁵Their bad debts are sold to agencies which specialize in debt recovery. Moreover, our lender offers standard debt contracts and the interest rate on their contracts does not differ across borrowers. Loan applications are either granted at the prevailing interest rate, or the request is turned down. Edelberg (2003) showed that systematically offering different interest rates to different agents, conditional on the borrowers observable characteristics (known as 'risk-based pricing' in the industry) is a comparatively recent phenomenon even in the US, and it is not currently practiced by Italian lenders.

and region of residence, seniority in the profession of the head and of the spouse, housing tenure, number of children, income and marital status. For each client, we have information on all current and past contracts and all applications (including rejected applications), for each of the three types of loans that are granted. Moreover, some of the loans, although authorized by the lender, were not activated by the customer.

For activated contracts, we have information on the type of contract, the amount financed or the credit limit, the amount repaid or the credit actually extended, and the currently outstanding debt and repayment status of the borrower. The data for accepted applicants is a cross sectional snapshot of all existing contracts, containing financial information on each contracts (including the price of the good, the item financed, the amount of credit extended, and the currently outstanding debt), as well as some demographics and other background information on the customer. For revolving credit, it includes the outstanding amount and the credit limit. The data also includes the bank's evaluation of the customer and the contract in terms of the repayment behavior of the borrower. More details on pricing and contracts can be found in Bertola, Hochguertel and Koeniger (2005) and Bicakova (2007). In about 15 percent of cases, applicants are refused credit. We have a random sample consisting of 5 percent of these rejected applications. This file records information similar to that which we have for accepted applicants. A few households have been rejected on some applications and have been successful in others.

To ensure our measure of default is as close as possible to actual default, we will concentrate on the installment credit contracts that are expired.⁶ Moreover, we select only customers for which we have complete

⁶Section 6.2 shows that the results are similar if we also include contracts which

information on the profession of household members, the province and region of residence, seniority in the profession of the head and of the spouse, housing tenure, number of children, income and marital status. This leaves us with almost 85,000 accepted and rejected applications.

Alessie, Hochguertel and Weber (2005) compare the data against a representative survey of the Italian population, the Survey of Household Income and Wealth. Compared to the Italian population as a whole, the households in our sample of credit applications are younger, are more likely to be living in the South, more likely to rent and have lower income than is typical of the Italian population. This is also documented in table 1, which focuses on both rejected and non-rejected applications in 1996 and 1998 and shows that Findomestic costumers are less likely to be self-employed, less likely to own their homes, and less likely to have a bank account.

An advantage of using administrative data is that measurement error is likely to be reduced compared to survey data. Since the lender uses this information when screening the applicant and when assessing whether action needs to be taken against borrowers in arrears, it has strong incentives to ensure its accuracy. However, administrative data contains less information about the characteristics of the household than most survey data, and any data on the household is self-reported by the applicant (who may have incentives to report falsely). Standard consistency checks on the data have been carried out as described in Alessie, Hochguertel and Weber (2005).

In order to construct our indicators of judicial efficiency and of borrowing from friends and relatives, we use two other data sources: the data on civil trials, provided by ISTAT, and the Survey of Household Income

have not come to the end of the agreed repayment period.

and Wealth (SHIW), a household level survey conducted by the Bank of Italy almost every second year. The first of these two data-sets is used to measure the quality of judicial enforcement, the second to construct an indicator that captures the availability of ‘non-market’ credit providers, such as relatives and friends. Judicial efficiency is proxied by the average length of trials in the civil courts in each Italian judicial district using data from the *Annuario di Statistiche Giudiziarie* for 1989-2000, published by ISTAT. The same proxy is used in Jappelli, Pagano and Bianco (2005).

Figure 1 plots the average length of civil trial (in years) in each region against the latitude of the city where the main court in the region is located: the regressions will also exploit variation over time. Fabbri and Padula (2004) document cross-regional differences in the time evolution of several proxies for the quality of judicial enforcement. In the figure there is a clear and significant geographical gradient, which shows that in Southern regions the average trial takes nearly twice as long as in the North. For example, Bolzano (which is near the Austrian border) has the most efficient court, and Catanzaro (in the far South) the least efficient. This shows that there are substantial differences between the regions: the effect of these differences will be studied in this paper.

To measure the availability of credit from informal sources, we exploit the Survey of Household Income and Wealth (SHIW), a representative survey of Italian households. This survey includes detailed questions about household characteristics, income, spending and assets. In particular, it records information on debts held by different lenders, such as banks, other financial institutions, and, importantly for us, informal credit from friends or non co-resident relatives.

We construct an indicator of the availability of informal credit mar-

kets by regressing a dummy for whether debt is held with relatives and friends on the region and the year in which the household lives at the moment the household applies for credit, as well as income and squared income, a dummy for couples, number of kids, a dummies for home-owners, for self-employed and for not holding a bank account. This yields a household level indicator of the availability of alternative credit providers and allows us to impute access to credit from informal sources (such as relatives and friends) in the lender's data. It measures the proportion of households with the given characteristics who report that they have some outstanding debt with friends, and/or family. This directly captures the availability of alternative credit sources, but the effect is uncertain. Recall that default will be more attractive if households can not be excluded from access to credit in the future. Thus if access to informal credit is more pervasive, then households have less incentive to repay debts incurred in the formal credit market. On the other hand, households may borrow from family and friends to help them meet payments on their formal credit contracts and therefore prevent default.

Figure 2 plots the distribution of the share of those borrowing from friends and family for each Italian region, listed from North to South. This is obtained by first computing our measure of reliance on family and friends financial help and showing for each region a box that is centered on the median and has left hinge equal to the 1st and right to the 3rd quartile of the distribution of our proxy for relatives' and friends' financial help. The graph shows that borrowing from friends and family is more prevalent in some Southern regions, such as Puglia, Calabria, and Sicilia, and that the prevalence is very limited in Sardina and Umbria. The graph also uncovers substantial variability within each Italian region, which suggests that the availability of informal sources of credit varies

both between and within regions.

The Findomestic data records whether households default on their debts. Figure 3 plots the proportion of loans that defaulted in each region against its latitude. It clearly shows that relatively fewer households repay their debts in the South than in the North. In the North, around 10 percent of households fail to repay their debts on schedule, but in Calabria and Sicilia over 20 percent fail to repay on schedule, which is twice that in Friuli and Veneto, for instance. These figures suggest a positive relationship between the repayment of debts and the ease with which borrowers can be punished if they default, which the rest of the paper explores in more detail.

4 The econometric model

Our aim is to investigate the effect of access to informal credit markets, and of judicial enforcement on borrower behaviour in the consumer credit market. We estimate upper and lower bounds of the effect of judicial enforcement and access to informal credit on repayment behaviour, and assess whether such households are more likely to apply for credit. To proceed, denote e as the repayment behaviour of the borrower, which takes the value 1 if the debt is repaid on schedule and zero if any scheduled repayment is missed. Denote by X a vector of observable characteristics that might affect the repayment behaviour. The vector X records the variables that the lender observes in making their credit granting decision and the additional variables that we have constructed. Finally, I_G is a binary variable that takes the value one if the credit request was granted, and zero if it was refused.

If the loan was granted we observe the repayment behaviour of the

borrower. Our exercise is to test how the household's characteristics affect whether the borrower defaults. However, we wish to deduce repayment behaviour accounting for the lending decision of our bank, which requires us to predict the likely repayment behaviour of households whose loan application was refused. How does the repayment behaviour of Italian households change with their characteristics? Obviously, to do this we need to assume that households who apply for loans from our lender are typical of all Italian applicants, e.g. given their observable characteristics, there is no difference between customers with our lender and customers who go elsewhere. This seems quite a strong assumption but can be consistent with economic theory. It will be true if, for instance, all lenders adopt the same strategy, conditional on the observable characteristics X . This seems plausible if all firms are profit maximizing in a competitive market (or if they all have the same market power). It does rule out lenders segmenting the market and adopting different strategies. For instance, in the US, 'sub-prime lenders' concentrate on low-income / high risk households that are denied credit by more traditional lenders. However, these type of sub-prime lenders did not operate in the Italian credit market at this time.

We also need to make an additional assumption if our results are representative of the whole Italian population: after accounting for the variables X , knowing that the agent has applied for a loan, does not predict whether it will be repaid. This is a much stronger assumption. It rules out that agents who are predictably bad borrowers are applying for credit.

4.1 The selection problem

We focus on the repayment behaviour of households and how repayment is related to household characteristics. For any household with characteristics X , their repayment probability is the sum of repayment if their credit application was granted multiplied by the probability of credit being granted, and repayment if it was refused multiplied by the probability of being refused credit.

$$E(e|X) = E(e|X, I_G = 1|X) Pr(I_G = 1|X) + E(e|X, I_G = 0|X) * Pr(I_G = 0|X) \quad (1)$$

We directly observe whether the household was granted credit, hence we know $Pr(I_G = 1|X)$, and since the household is either granted or refused credit, we also know that $Pr(I_G = 0|X) = 1 - Pr(I_G = 1|X)$. Since the household's repayment behaviour is observed if it was granted credit, we can also construct the sample analog of $E(e|X, I_G = 1)$. However, the repayment behaviour of households refused credit can not be directly observed hence we can not construct $E(e|X, I_G = 0)$. This is the selection problem: some way must be found to estimate $E(e|X, I_G = 0)$ from what is observed, although we know it must lie between zero and one. This selection issue is crucial if one wants to evaluate the determinants of default. For instance, observing that default decreases with income does not necessarily mean that poorer households are more likely to default if they also borrow less often. Similarly, if one is interested on the time evolution of default, failing to account for the changing composition of the pool of borrowers can be misleading about the true dynamics of default.

Notice that if rejected applicants are less likely to repay (implying

that lenders screen out bad risks), the sample analog of $E(e|X, I_G = 1)$ overestimates $E(e|X)$. While there are several standard econometric techniques to handle this type of selection problem, each imposes economic assumptions. We now discuss the economic assumptions that are needed for identification.

4.1.1 Selection by observables, i.e. “Rubin”

In matching models, there is some set of variables, W , by which observations are sorted. Households with the same W are matched: households for which $I_G = 1$ replace those households, with the same W , for which $I_G = 0$. Formally, this requires that:

$$e \perp I_G | W$$

which implies that $Pr(I_G = 1|X, W, e) = Pr(I_G = 1|X, W)$.⁷ In our framework, this amounts to requiring that the set W affects the credit scoring algorithm, but not repayment behaviour. This is unlikely to be satisfied by the data if the lender is rational and maximizes profits. Why would the lender use a variable in their screening procedure if it did not affect repayment behaviour? Hence this assumption seems inconsistent with lender rationality. However, it will be satisfied, if, for instance, the lender discriminates for non-economic reasons against subgroups in society (such as ethnic minorities) but this would be inconsistent with profit maximization.

4.1.2 Selection by unobservables, i.e. “Heckman”

A second popular method of solving the selection problem is through specifically modelling the selection process. Identification using this

⁷Matching also requires the common support assumption to be satisfied. This implies that $0 < Pr(I_G = 1|X, W) < 1$.

method requires some exclusion restriction (except in the case of identification via functional form assumptions). Formally, identification requires that for some set of variables W we have:

$$\text{(a)} \quad E(e|X) = E(e|X, W)$$

$$\text{(b)} \quad Pr(I_G = 1|W, X) \neq Pr(I_G = 1|X)$$

meaning that the W affects whether the application is rejected but does not affect repayment. Assumptions (a) and (b) are not very attractive. If W enters the screening procedure, then as before, rationality implies that it is likely to affect repayment behaviour.

4.1.3 The bounds

Neither selection by unobservables nor matching seem to be particularly attractive with the data at hand. Both impose economic assumptions that are difficult to reconcile with lender rationality: lenders are likely to include variables in their screening procedure only if they predict likely repayment behaviour. However, if we make weaker assumptions, we can place bounds on the estimated effects of interest. Recall that we have modelled whether the household repays, e , as a binary variable equal to one if the debt is repaid on schedule, and zero if not. We can re-write equation (1) as:

$$\begin{aligned} Pr(e = 1|X) &= Pr(e = 1|X, I_G = 1) * Pr(I_G = 1|X) \\ &\quad + Pr(e = 1|X, I_G = 0) * Pr(I_G = 0|X) \end{aligned}$$

Recall that the data do not allow us to identify $Pr(e = 1|X, I_G = 0)$: we cannot observe whether households repay (or would have repaid) their loan if they were never granted credit. However, we know that the probability of repaying lies between zero and one, thus we can proceed as

in Manski (1989), and place upper and lower bounds on the effect of the variables of interest. The lower bound assumes a zero probability of repaying for those who are refused credit, while the upper that this probability is one. This gives:

$$K_{0X} = Pr(e = 1|X, I_G = 1) * Pr(I_G = 1|X) \leq Pr(e = 1|X) \leq Pr(e = 1|X, I_G = 1) * Pr(I_G = 1|X) + Pr(I_G = 0|X) = K_{1X} \quad (2)$$

That is, we define the lower bound as K_{0X} and the upper bound as K_{1X} . These bounds can be identified since we have data on rejected applications, from which we can construct the probability that a credit application was granted, $Pr(I_G = 1|X)$, and since we observe the repayment behaviour of households granted credit we can construct $Pr(e = 1|X, I_G = 1)$.

Suppose now that we want to measure the effect of access to informal credit markets, measured by the incidence of financial help from friends and family, on default rates. If one had point identification, the effect of family financial help is captured by the relevant β in the estimation of $Pr(e = 1|X\beta)$. Otherwise, we could split the sample by whether the reliance on family financial help is high (above median) or low. Defining $Z = H$ if the reliance is high and $Z = L$ if it is low, the difference in effort across high and low reliance is:

$$Pr(e = 1|\tilde{X}, Z = L) - Pr(e = 1|\tilde{X}, Z = H)$$

where we partition X into \tilde{X} and Z . This probability is bounded between

$K_{1XL} - K_{0XH}$ and $K_{0XL} - K_{1XH}$, where:

$$\begin{aligned}
K_{0XL} &= Pr(e = 1|\tilde{X}, I_G = 1, Z = L) * Pr(I_G = 1|\tilde{X}, Z = L) \\
K_{1XL} &= Pr(e = 1|\tilde{X}, I_G = 1, Z = L) * Pr(I_G = 1|\tilde{X}, Z = L) + \\
&\quad + \left[1 - Pr(I_G = 1|\tilde{X}, Z = L)\right] \\
K_{0XH} &= Pr(e = 1|\tilde{X}, I_G = 1, Z = H) * Pr(I_G = 1|\tilde{X}, Z = H) \\
K_{1XH} &= Pr(e = 1|\tilde{X}, I_G = 1, Z = H) * Pr(I_G = 1|\tilde{X}, Z = H) + \\
&\quad + \left[1 - Pr(I_G = 1|\tilde{X}, Z = H)\right]
\end{aligned}$$

The intuition is simple. Suppose that the effect of Z on repayment is positive, then the smallest difference between a high value $Z = H$ and a low value $Z = L$ for access to credit from family and friends Z is given by $K_{1XL} - K_{0XH}$. Here, K_{1XL} represents the highest possible value for when $Z = L$ while K_{0XH} represents the lowest possible value for when $Z = H$. In contrast the largest difference is given by $K_{0XL} - K_{1XH}$, the difference between the lowest possible value when $Z = L$ and the highest possible value when $Z = H$. Hence the difference in the estimated effect of when Z is high, and when Z is low, must be between $K_{1XL} - K_{0XH}$ and $K_{0XL} - K_{1XH}$. If, instead, the effect of Z is negative, $K_{1XH} - K_{0XL}$ becomes the lower bound, and $K_{0XH} - K_{1XL}$ the upper bound.

4.1.4 Tightening the bounds

Suitable assumptions tighten the bounds, which could otherwise be large. Lenders have strong incentives to refuse credit to households which are less likely to repay their debts, hence households refused credit are likely to be worse credit risks. If this were not true then it would imply lenders were rejecting low risk and accepting high risk applicants, which would be inconsistent with their motive to screen customers, and with profit

maximization. If rejected applicants are weakly less likely to repay their debts than households whose application was accepted then:

$$Pr(e = 1|X, I_G = 0) \leq Pr(e = 1|X, I_G = 1)$$

Using this inequality means that equation (2) becomes:

$$Pr(e = 1|X, I_G = 1) * Pr(I_G = 1|X) \leq Pr(e = 1|X) \leq Pr(e = 1|X, I_G = 1) = \tilde{K}_{1X} \quad (3)$$

which narrows the bounds, since the upper boundary, \tilde{K}_{1X} , is now tighter. With these narrower bounds the effect of informal credit markets, measured by the extent of lending through friends and family, is instead bounded between $\tilde{K}_{1XL} - K_{0XH}$ and $K_{0XL} - \tilde{K}_{1XH}$ where:

$$\begin{aligned} \tilde{K}_{1XL} &= Pr(e = 1|X, I_G = 1, Z = L) \\ \tilde{K}_{1XH} &= Pr(e = 1|X, I_G = 1, Z = H) \end{aligned}$$

5 Estimation

In order to estimate the bounds one needs the sample analog of $Pr(e = 1|X, I_G = 1)$ and $Pr(I_G = 1|X)$. We will use both a fully parametric estimator and a semi-parametric estimator to construct these probabilities. Calculating the lower bound means estimating:

$$Pr(e = 1|X, I_G = 1) Pr(I_G = 1|X) \quad (4)$$

which is equivalent to estimating:

$$E(e \cdot I_G|X) \quad (5)$$

That is, estimating the lower boundary is equivalent to estimating the proportion of households who are both granted credit and repay that credit on time. While for the upper bound estimating:

$$Pr(e = 1|X, I_G = 1) Pr(I_G = 1|X) + Pr(I_G = 0|X) \quad (6)$$

is equivalent to estimating:

$$E(e \cdot I_G - I_G | X) + 1 = 1 - E(I_G(1 - e) | X) \quad (7)$$

That is, estimating the lower boundary is equivalent to estimating one minus the proportion of households who are both granted credit and do not repay that credit on time. Lastly, the tightened upper bound is the probability that a household which is given credit repays the debt. This is the same as a naive estimate (which ignores selection) of the effect of the variable of interest on repayment behaviour. The upper, tightened upper and lower boundaries are easily calculated.

The parametric estimation of the lower, upper and tightened upper bound is straightforward. We use standard probit regressions to provide the first set of results. Additionally, we estimate the lower, upper and tightened upper bound by using a semi-parametric approach. Probit regression is a maximum likelihood estimator which maximizes the function:

$$\prod_i F(X_i\beta)^{d_i} [1 - F(X_i\beta)]^{1-d_i}$$

where d_i is a dummy variable which takes the value one if the event happens (for example, that both credit is granted and the debt is repaid on schedule for the lower bound), and zero otherwise. Probit regressions make parametric assumptions about both the functional form of the index function $X_i\beta$ and the distribution of the error term, assuming that F is the normal distribution.

In order to free our estimates from distributional assumptions, we also employ a semi-parametric estimator. The tighten upper bound to the probability of repayment, $Pr(e = 1 | X, I_G = 1)$, is estimated by minimizing the sample analog of:

$$\|e - E(e|I_G = 1, X)\|$$

where we approximate $E(e|I_G = 1, X)$ with a fourth order polynomial in a linear index, $X\beta$.⁸ This procedure is similar to that suggested by Ichimura (1993), with two important differences. First, Ichimura employs kernel methods as approximation method while we favour series due to the large dimensionality of our problem. Second, while we are only interested in the probabilities, Ichimura's focus is on the estimation of β , which represents a more difficult task. The probability of the application not being rejected is estimated in a similar way and the lower and upper bounds are computed using (4) and (6), respectively.

Standard errors are obtained by a two stages bootstrap: in the first we bootstrap the sample used to compute our indicator of relatives' and friends' financial help, in the second we bootstrap the sample drawn from our lender archives.

6 Results

Table 2 presents estimates of the bounds to the probability of repayment. We concentrate on installment credit contracts, although results are similar for revolving credit contracts and personal loans. The columns report the lower and the upper bound, as well as the tightened upper bound, which is the naive estimate without controlling for selection. All regressions feature a time trend and agency dummies, which implies that our estimates are not biased by unobservable time and geographic effects.⁹

⁸Trying higher order polynomials does not affect the results. Consistency relies on the correlation between the higher order terms and the error, defined as $e - E(e|I_G = 1, X)$, to vanish as the sample size grows.

⁹Agency dummies are defined on the basis of the province where the agency dealing with the contract is located. Italy is divided into 20 regions and about 100 provinces.

The results in the first column show that the lower bound to the probability of repayment is an increasing and concave function of job seniority and income. This is consistent with the intuitive notion that wealthier households are less likely to default, though at a decreasing pace. The lower bound is higher for home-owners and lower for self-employed. The former are typically stable income earners, while for the latter, income is typically more volatile and risky, which makes default more likely.

Our measure of the quality of judicial enforcement, the average length of civil trials, reduces the lower bound to the probability of repaying. The proportion of households which repay their loans decreases significantly as reliance on friends and family for financial help increases. The degree of competition in the credit market, measured by the number of bank branches per bank in the province, is instead positively correlated with the proportion of households who repay. The effect of access to credit from family and friends accords with the idea that those who have better outside options are more likely to default. Even if they are permanently excluded from the formal credit market, they can still borrow from friends and family in the informal credit market.

The results in the second column of table 2 refer to the upper bound to the probability of repayment. The estimated coefficients have the same sign, except for the degree of competition, measured by the number of branches per bank in the province, which switches from positive to negative. We interpret this result as reflecting that higher competition causes lenders to weaken credit standards, which reduces the average quality of the borrower and raises *ex post* default rates. The upper bound is an increasing and concave function of both job tenure and income, is higher among home-owners and lower among self-employed. Judicial

enforcement and reliance on family and friends financial help reduce the upper bound.

The last column refers to the tightened upper bound, which is the probability of repaying among those actually given credit. It is also the estimate had we ignored selection and focussed on the sample of non-rejected applications. The results imply that the probability of repaying is an increasing and concave function of income and job-tenure, is higher among home-owner, lower among self-employed and those who do not hold a bank account, and is negatively related to length of trials, financial help from family and friends. If selection is not an issue, the changes in the probability of repayment can be inferred from the coefficients shown in the last column of table 2, which focuses on the probability of repayment among non-rejected applicants. However, the selection issue arises as the sample of rejected applications is likely to be different from the applications which are not rejected.

In order to appreciate the importance of selection, we explore how repayment changes if, say, the value of the index of reliance on friends' and relatives' financial help increases from the minimum to the maximum value of its distribution. If there is a selection problem (refused and accepted credit applications are systematically different) then the upper and lower bounds to the probability of repayment can still be constructed as described in section 4. Two situations may arise: either the lower and upper bounds to the changes in the repayment probability have the same sign, (both negative or both positive) or they have different signs. In the former instance, we can identify the sign of the effect of the variable of interest (since the true effect is somewhere between the upper and lower bound), in the latter case we can not. The results are reported in table 3 and refer to home-ownership, job tenure, income, length of trials, number

of branches and our indicator of reliance on family and friends financial help.

The first row of table 3 focuses on the effect of home-ownership and is obtained by setting the home-owner dummy to 0 and 1 in turn, evaluating the lower, the upper and the tighten upper bounds for the two values of the home-owner dummy and bounding the amount by which the probability of repayment changes with the home-ownership status as described in section 4.1.4. For presentation purposes, the bounds are then averaged across the sample. The other rows of the table are obtained in a similar way.

The first column of the table 3 shows how much the probability of repayment changes with home-ownership if one ignores selection. Ignoring selection, the global effect of home-ownership on repayment implies that home-owners are 1.6% more likely to repay than non-home owners. The other columns of the table allow for non-random selection and show that switching the home-owner dummy from 0 to 1 causes the change in the probability of repaying to be bounded between -0.085 (with standard error equal to 0.003) and 0.143 (0.003) if one uses the upper bound, and from -0.061 (0.003) and 0.128 (0.003) if one uses the tightened upper bound. Hence, one cannot rule out that the effect of home-ownership on repayment is zero. Therefore, neglecting the selection problem can lead to unwarranted conclusions. The lower bound increases with home-ownership, a likely consequence of the loan selection process, but not enough: the lower bound for home-owners is smaller than the upper bound for non home-owners, everything else equal. The effect of home-ownership on the probability of repayment can be visualized from figure 4. The figure plots the lower (solid line), upper (dotted) and tighten upper (dashed) bounds of the probability of repayment as function of home-

ownership status. The figure clarifies why we cannot rule out zero effect: one could draw a straight line without crossing any of the bounds.

The global effect of raising both income and job tenure is negative if one ignores selection (respectively, -0.421 and -0.374). Accounting for selection implies the effect of job tenure is bounded between -0.808 (0.014) and -0.283 (0.137) and that of income between -0.785 (0.200) and -0.245 (0.346). Whether or not one ignores selection, the effect of job tenure and income on repayment is negative. This is due to the probability of repayment not being a monotone function of job tenure and income, as shown in figures 5 and 6, which plot the bounds to the probability of repaying against, respectively, the distribution of job tenure and income and imply that the probability of repayment is a concave function of job tenure and income.

The sign of the effect of length of trial is not identified, as shown in figure 7, which plots the bounds to the probability of repayment against length of trial. The effect is bounded between -0.182 (0.010) and 0.054 (0.009) using the wide bounds and between -0.166 (0.010) and 0.026 (0.010) using the tightened bounds. Since zero lies between the upper and the lower bounds we cannot rule out that there is no effect. The results on enforcement are not entirely surprising, since consumer credit is unsecured and collateral is the main channel through which the quality of enforcement affects borrowers' behaviour.¹⁰

Increasing the number of branches per banks in the province reduces repayment. The change in the probability of repayment are bounded from below and from above by a negative number: by between -0.168 (0.032) and -0.029 (0.034) or -0.146 (0.033) and -0.019 (0.035) using the tighten bounds. However, since the upper bounds to the changes are not

¹⁰This does not rule out that there is an effect of enforcement on loan selection.

very precisely estimated, the results do not rule out either a positive or a negative effect of banking concentration on repayment.

Increasing access to credit from friends and family reduces the probability of repayment by between -0.459 (0.039) and -0.015 (0.029) or by 0.454 (0.04) and -0.121 (0.039) using the tightened bounds. It supports the hypothesis that the availability of informal credit weakens incentives to repay. This is also clear from figure 8, which plots the lower, upper and tighten upper bounds to the probability of repaying as function of relatives' and friends' financial help. We now turn to the semi-parametric results.

6.1 Semi-parametric results

The semi-parametric results are broadly consistent with the parametric. However, for job tenure and for income, abandoning the parametric assumptions makes identification harder, as shown in figure 9 and 10. Figure 11 shows that the upper and the tightened upper bounds increase with the length of trials, while the lower bound is flat. This implies that one cannot exclude that judicial enforcement does not affect the probability of repaying. The effect of borrowing from relatives and friends on the probability of repaying is shown in figure 12. The figure shows that the lower, the upper and the tightened upper bounds decrease as financial help from family and friends increases. The decrease is sizable and statistically significant, as shown by the diamonds, the plus, and the stars placed two standard deviations above and below the lower, the upper and the tighten upper bound.

As before, we quantify the effect on the probability moving from the minimum to the maximum of each of the variables of interest. The effect of home-owning on probability of repaying is between -0.111 (0.007) and

0.116 (0.007), or -0.088 (0.007) and 0.098 (0.008) using the tightened bounds. The effects of job tenure and income is not identified. Increasing job-tenure causes the probability of repayment to change by between -0.161 (0.029) and 0.150 (0.005) or by between -0.135 (0.005) and 0.121 (0.004) if using the tightened bounds; for increasing income, the probability of repayment is bounded between -0.178 (0.025) and 0.131 (0.012) or by -0.153 (0.025) and 0.098 (0.016).

The effect of the degree of banks competition on repayment is negative and is bounded between -0.653 (0.031) and -0.180 (0.029) if one uses the tightened bounds. The effect of judicial enforcement is instead not identified: the changes of the probability of repayment are bounded by a negative number from below and a positive number from above. For access to credit from family and friends, the probability of defaulting decreases by between -0.576 (0.044) and -0.322 (0.072) on the tightened upper bound.

6.2 Robustness checks

Our definition of default focuses on expired contracts only and neglects those contracts that are still alive in 1999, the date at which the data are drawn from our lender archive. Therefore, compared to the whole population of contracts, the selected contracts are on average older, which might be source of concern if the composition of costumers changes over time. We then extend our definition of default to include also not expired contracts and estimate the lower, the upper and the tighten upper bounds both parametrically and semi-parametrically.

The parametric results are reported in table 5 and are consistent with the results obtained on the expired contracts only. The lower, the upper and the tighten upper bounds are larger for home-owners, smaller for

self-employed and for those who do not hold a bank account. Moreover, the bounds to the probability of repaying are increasing and concave in job tenure and income, and decrease with the length of trials and with the reliance on relatives' and friends' financial help. Table 6 quantifies the effect on the probability of repayment of changing home-ownership status, job tenure, income, degree of banks competition, length of trials and reliance on relatives' and friends' financial help. The results confirms that the effect of financial help is negative and do not rule out a zero effect of judicial enforcement. The same message is conveyed by the semi-parametric results. Figure 13 and 14 plot the lower, upper and tighten upper bounds to the probability of repayment as function of judicial enforcement and financial help respectively. Again, the former effect is not identified, but the latter is clearly negative.

As further check, we split our sample between Northern and Southern Italian provinces. If our results on family and friends financial help are solely driven by the North-South divide, one should see the effect of financial help from relatives and friends to disappear when focussing on the North or on the South separately. Tables 7 and 8 report the parametric estimates for Northern and Southern provinces and reveal that the splitting the sample between North and South does not change the overall picture. Both for Northern and Southern provinces, the lower, upper and tightened upper bounds decrease with the reliance on relatives' and friends' financial help. Figures 15 and 16 focus on the semi-parametric estimates instead, confirms that the lower upper and tighten upper bounds decreases with reliance on family and friends financial help both for Northern and Southern provinces, but imply that the effect is identified only for Southern provinces. This means that the difference between markets and institutions between Northern and Southern provinces

is part of the story, but cannot explain fully our results.

7 Conclusions

Using a leading Italian lender's administrative data on credit applications, we are able to assess how features of the market affect repayment behaviour. Identifying the factors affecting repayment is not trivial. A selection issue arises because we do not observe the repayment behaviour of those households that were refused credit. Two popular methods for addressing selection require imposing the economic restriction that the lender's screening procedure is unrelated to the potential borrowers repayment behaviour, which is unlikely to be satisfied. At the cost of losing point identification, we impose less stringent assumptions and provide upper and lower bounds of the likely effect of the variables of interest. Moreover, we show that plausible economic assumptions (that loans are made to less risky households) can reduce the size of these bounds.

In this study we investigate the effect of judicial institutions, of informal credit and of bank competition on repayment. We measure judicial enforcement with the average length of civil trials, while we use credit from friends and family to measure the availability of informal sources of credit. We also measure competition using the total number of bank branches in each province. Our results show that the quality of judicial enforcement makes little difference to whether the borrower defaults or repays his loan. In contrast, access to credit from friends and family has large effects (the bounds exclude zero from the confidence interval). The effect of informal credit markets on whether the debt was repaid is both economically and statistically significant. This large effect is consistent with theory. Recall that both Kehoe and Levine (1993) and Kocherlakota

(1996) argued that incentives to repay depended on the punishment for default. For the small unsecured loans that our lender specializes in, the judicial process is relatively unimportant since it is rarely invoked. This is perhaps unsurprising given the small size of the typical loan that we examine, and the fact that these loans are uncollateralized. The effect of judicial enforcement is economically small (one tenth of the effect of family and friends) and statistically insignificant. Instead, defaulters are punished by being denied further loans. But this means that borrowing from family and friends improves the household's outside option and reduces the penalty for default. Households with access to these informal credit markets view exclusion from the formal credit market as less onerous since they can still borrow from friends and family should the need arrive. Thus, if households have alternative credit sources (through informal credit markets provided by family and friends) then their incentives to repay debts in the formal market are much lower.

These results - that access to informal sources of credit increase the default rate in the formal sector - suggests that there is moral hazard, in the sense that access to credit from family and friends reduces repayment in the formal sector. However, unlike conventional moral hazard where the loan from the lender changes repayment behaviour on that loan, we instead show that access to alternative credit sources changes repayment behaviour on the loan granted by the lender.

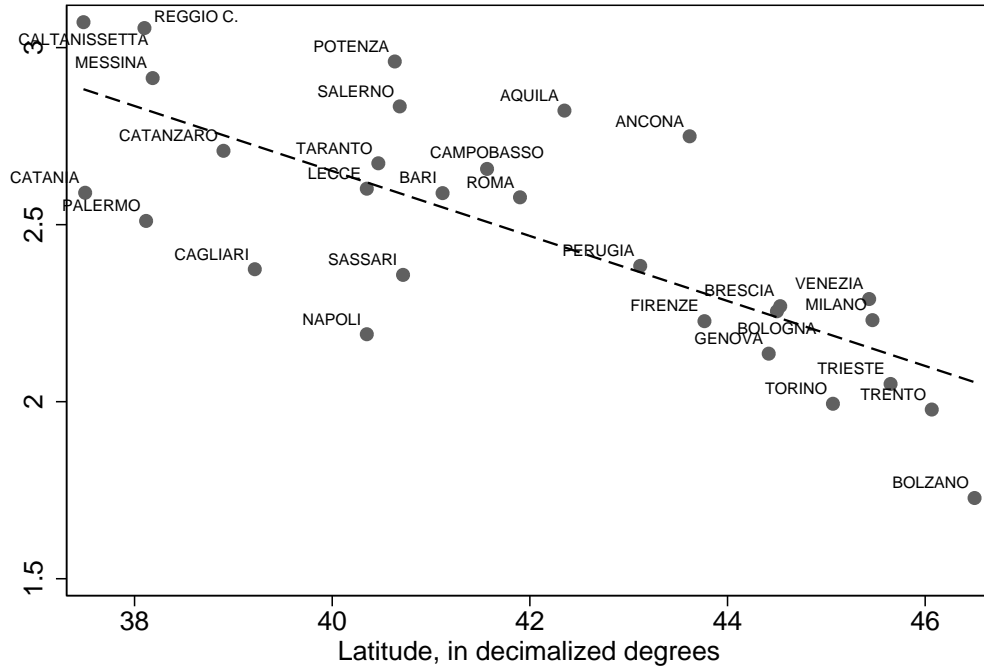
References

- Alessie, Rob, Hochguertel, Stefan and Guglielmo Weber, (2005), "Consumer Credit: Evidence from Italian Micro Data," *Journal of the European Economic Association*, 3, pp. 144-178.
- Banerjee Abhijit, and Newman, Andrew, (1998) "Information, the Duel Economy, and Development," *Review of Economic Studies*, 65, pp. 631-653.
- Bertola, Giuseppe, Hochguertel, Stefan and Winfried Koeniger, (2005), "Dealer Pricing of Consumer Credit," *International Economic Review*, 46, pp. 1103-1142.
- Bicakova, Alena, (2007) "Does the Good Matter? Evidence on Moral Hazard and Adverse Selection from Consumer Credit Market", *Giornale degli Economisti and Annali di Economia*, forthcoming.
- Casolaro, Luca, Gambacorta Leonardo and Luigi Guiso, (2005), "Regulation, Formal and Informal Enforcement and the Development of Household Loan Market: lessons from Italy", in *The Economics of Consumer Credit*, ed. by Giuseppe Bertola, Richard Disney and Charles Grant, MIT press, Cambridge, MA.
- Edelberg, Wendy, (2003), "Risk-based pricing of interest rates in household loan markets," *FEDS Working Paper No. 2003-62*.
- Edelberg, Wendy, (2004) "Testing for Adverse Selection and Moral Hazard in Consumer Loan Markets" *FEDS Working Paper No. 2004-09*.
- Fabbri, Daniela and Mario Padula, (2004), "Does Poor Legal Enforcement Make Households Credit-Constrained?," *Journal of Banking and Finance*, 28, pp. 2369-2397.

- Fay, Scott, Hurst, Erik, and Michelle White (2002), "The Household Bankruptcy Decision", *American Economic Review*, 92, pp. 706-718.
- [1] Ghosh, Parikshi, Mookherjee, Dilip and Debray Ray, (2000), "Credit rationing in developing countries: an overview of the theory," in *A Reader in Development Economics*, edited by Dilip Mookherjee and Debraj Ray, London: Blackwell
- Jaffee, Dwight, M., and Tom Russel, (1976), "Imperfect information, uncertainty, and credit rationing," *Quarterly Journal of Economics*, 80, pp. 651-666.
- Jappelli, Tullio, Pagano Marco and Magda Bianco, (2005), "Courts and Banks: the Effect of Judicial Enforcement in Credit Markets," *Journal of Money Credit and Banking*, 37, pp. 223-245.
- Ichimura, Hideiko, (1993), "Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single-Index Models," *Journal of Econometrics*, 58, pp. 71-120.
- Kehoe, Timothy J., and David K. Levine, (1993), "Debt-Constrained Asset Markets," *Review of Economic Studies*, 60, pp. 865-888.
- Kocherlakota, Narayana, (1996), "Implications of efficient risk sharing without commitment", *Review of Economic Studies*, 63, pp. 595-609.
- Manski, Charles, (1989), "The anatomy of selection problem," *Journal of Human Resources*, 24, pp. 343-360.
- Manski, Charles, (1990), "Nonparametric bounds on treatment effects," *American Economic Review*, 80, pp. 319-323.

Stiglitz, John E., and Andrew Weiss, (1981), "Credit rationing in markets with imperfect information," *American Economic Review*, 71, pp. 393-410.

Figure 1: Length of Trial and Latitude



Note. The length of trial refers to the average length of civil trials in the main city in each Italian judicial district, measured in years. Data constructed from the *Annuario di Statistiche Giudiziarie* for 1989-2000, published by ISTAT.

Table 1: Summary statistics

	SHIW		Findomestic All contracts		Findomestic Not-rejected contracts	
	1995	1998	1996	1998	1996	1998
North	48.63	48.02	32.55	32.32	33.79	31.84
Central	18.25	19.06	24.77	22.29	26.96	22.92
South	33.22	32.92	42.68	45.39	39.25	45.23
Home-owner	65.66	66.6	41.87	43.62	47.10	46.37
Income	25,510	27,337	14,498	14,641	15,360	15,050
Self-employed	16.36	16.88	13.98	12.85	13.01	12.83
No bank account	31.41	27.2	55.49	60.85	49.84	59.30
N. of obs.	8,135	7,147	10,645	28,998	7,700	24,802

Note. The statistics from SHIW are computed using sample weights. For the Findomestic sample, rejected and not-rejected contracts for expired contracts are included in the third and fourth column; the fifth and the sixth columns focus on the not-rejected contracts only. Income is expressed in 2000 prices. The number of observations is maximum number of non-missing observations across the variables summarized in the table.

Table 2: Parametric estimates of the bounds on the probability of not-defaulting.

	Lower Bound	Upper Bound	Tightened Upper Bound
Home-owner	0.166 (0.012)***	0.033 (0.013)***	0.069 (0.014)***
Job tenure	0.386 (0.031)***	0.071 (0.016)***	0.167 (0.019)***
Squared job tenure	-0.081 (0.011)***	-0.010 (0.005)***	-0.032 (0.006)***
Income	0.129 (0.017)***	0.047 (0.014)***	0.072 (0.011)***
Income-squared	-0.003 (0.003)	-0.001 (0.002)	-0.001 (0.001)
Length of trials	-0.048 (0.008)***	-0.063 (0.008)***	-0.070 (0.009)***
Number of branches	-0.000 (0.000)	-0.003 (0.000)	-0.002 (0.000)
Self-employed	-0.199 (0.015)***	-0.164 (0.016)***	-0.190 (0.017)***
No bank account	-0.301 (0.017)***	-0.298 (0.015)***	-0.323 (0.016)***
Reliance on friends financial help	-2.362 (0.249)***	-0.693 (0.232)***	-1.354 (0.254)***
Observations	84,613	84,613	74,685

Note. Bootstrap standard errors in parenthesis. * means significant at 5 percent level, ** significant at 1 percent level, *** significant at 0.1 percent level. The regression also included a full set of provincial and year dummies. Income is in 10,000 Euros.

Table 3: Changes in the probability of defaulting: parametric estimates

	No selection	$K_{1XL} - K_{0XH}$	$K_{0XL} - K_{1XH}$	$\tilde{K}_{1XL} - K_{0XH}$	$K_{0XL} - \tilde{K}_{1XH}$	Mean
Home-owner	0.016 (0.003)***	-0.085 (0.003)***	0.143 (0.003)***	-0.061 (0.003)***	0.128 (0.003)***	0.445
Job tenure	-0.421 (0.138)**	-0.841 (0.014)***	0.120 (0.093)	-0.808 (0.014)***	-0.283 (0.137)*	0.820
Income	-0.374 (0.347)	-0.815 (0.200)***	-0.050 (0.305)	-0.785 (0.200)***	-0.245 (0.346)	1.468
Length of trials	-0.075 (0.010)***	-0.182 (0.010)***	0.054 (0.009)***	-0.166 (0.010)***	0.026 (0.010)**	3.023
Number of branches	-0.129 (0.034)***	-0.168 (0.032)***	-0.029 (0.034)	-0.146 (0.033)***	-0.019 (0.035)	50.289
Reliance on friends	-0.183 (0.039)***	-0.459 (0.039)***	-0.015 (0.029)	-0.454 (0.040)***	-0.121 (0.039)**	0.099

Note. Bootstrap standard errors in parenthesis. * means significant at 5 percent level, ** significant at 1 percent level, *** significant at 0.1 percent level. The last column reports the means of the variables given in left-hand column. The length of trials is expressed in years.

Table 4: Changes in the probability of defaulting: semi-parametric estimates

	$K_{1XL} - K_{0XH}$	$K_{0XL} - K_{1XH}$	$\tilde{K}_{1XL} - K_{0XH}$	$K_{0XL} - \tilde{K}_{1XH}$
Home-owner	-0.111 (0.007)***	0.116 (0.007)***	-0.088 (0.007)***	0.098 (0.008)***
Job tenure	-0.161 (0.029)***	0.150 (0.005)***	-0.135 (0.029)***	0.121 (0.004)***
Income	-0.178 (0.025)***	0.131 (0.012)***	-0.153 (0.025)***	0.098 (0.016)***
Length of trials	-0.122 (0.012)***	0.132 (0.012)***	-0.092 (0.012)***	0.112 (0.013)***
Number of branches	-0.703 (0.034)***	0.278 (0.050)***	-0.653 (0.031)***	-0.180 (0.029)***
Reliance on friends	-0.608 (0.046)***	0.098 (0.096)	-0.576 (0.044)***	-0.322 (0.072)***

Note. Bootstrap standard errors in parenthesis. * means significant at 5 percent level, ** significant at 1 percent level, *** significant at 0.1 percent level. The last column reports the means of the variables given in left-hand column. The length of trials is expressed in years.

Table 5: Parametric estimates of the bounds on the probability of not-defaulting: expired and not-expired contracts.

	Lower Bound	Upper Bound	Tightened Upper Bound
Home-owner	0.164 (0.009)***	0.054 (0.009)***	0.082 (0.009)***
Job tenure	0.357 (0.021)***	0.103 (0.012)***	0.172 (0.013)***
Squared job tenure	-0.074 (0.008)***	-0.017 (0.003)***	-0.032 (0.004)***
Income	0.119 (0.013)***	0.057 (0.010)***	0.075 (0.009)***
Income-squared	-0.002 (0.003)	-0.001 (0.001)	-0.002 (0.001)
Length of trials	-0.019 (0.010)*	-0.032 (0.007)*	-0.039 (0.008)*
Number of branches	-0.000 (0.001)	-0.002 (0.000)	-0.001 (0.000)
Self-employed	-0.189 (0.016)***	-0.165 (0.016)***	-0.183 (0.017)***
No bank account	-0.277 (0.016)***	-0.275 (0.012)***	-0.291 (0.013)***
Reliance on friends financial help	-2.189 (0.260)***	-0.791 (0.202)***	-1.312 (0.231)***
Observations	114,257	114,257	104,246

Note. Bootstrap standard errors in parenthesis. * means significant at 5 percent level, ** significant at 1 percent level, *** significant at 0.1 percent level. The regression also included a full set of provincial and year dummies. Income is in 10,000 Euros.

Table 6: Changes in the probability of defaulting: parametric estimates, expired and not-expired contracts.

	No selection	$K_{1XL} - K_{0XH}$	$K_{0XL} - K_{1XH}$	$\tilde{K}_{1XL} - K_{0XH}$	$K_{0XL} - \tilde{K}_{1XH}$
Home-owner	0.018 (0.002)***	-0.057 (0.002)***	0.114 (0.002)***	-0.041 (0.002)***	0.104 (0.002)***
Job tenure	-0.414 (0.086)***	-0.843 (0.011)***	-0.010 (0.064)	-0.820 (0.011)***	-0.307 (0.085)***
Income	-0.439 (0.309)	-0.701 (0.218)**	-0.265 (0.308)	-0.681 (0.218)**	-0.341 (0.307)
Length of trials	-0.038 (0.008)***	-0.114 (0.010)***	0.061 (0.007)***	-0.103 (0.010)***	0.042 (0.009)***
Number of branches	-0.083 (0.033)*	-0.117 (0.050)*	-0.011 (0.027)	-0.100 (0.052)	0.002 (0.037)
Reliance on friends	-0.168 (0.034)***	-0.395 (0.041)***	-0.042 (0.026)	-0.393 (0.043)***	-0.121 (0.035)***

Note. Bootstrap standard errors in parenthesis. * means significant at 5 percent level, ** significant at 1 percent level, *** significant at 0.1 percent level. The last column reports the means of the variables given in left-hand column. The length of trials is expressed in years.

Table 7: Parametric estimates of the bounds on the probability of not-defaulting, North.

	Lower Bound	Upper Bound	Tightened Upper Bound
Home-owner	0.245 (0.016)***	0.081 (0.018)***	0.121 (0.019)***
Job tenure	0.492 (0.049)***	0.105 (0.023)***	0.203 (0.028)***
Squared job tenure	-0.105 (0.018)***	-0.016 (0.008)***	-0.039 (0.010)***
Income	0.134 (0.036)***	0.070 (0.019)***	0.092 (0.020)***
Income-squared	-0.006 (0.008)	-0.002 (0.003)	-0.003 (0.004)
Length of trials	-0.015 (0.010)	-0.039 (0.013)	-0.038 (0.013)
Number of branches	-0.001 (0.002)	-0.007 (0.002)	-0.006 (0.002)
Self-employed	-0.224 (0.019)***	-0.187 (0.025)***	-0.211 (0.026)***
No bank account	-0.199 (0.022)***	-0.165 (0.024)***	-0.185 (0.026)***
Reliance on friends financial help	-2.801 (0.372)***	-1.145 (0.433)***	-1.791 (0.463)***
Observations	46,186	46,186	41,085

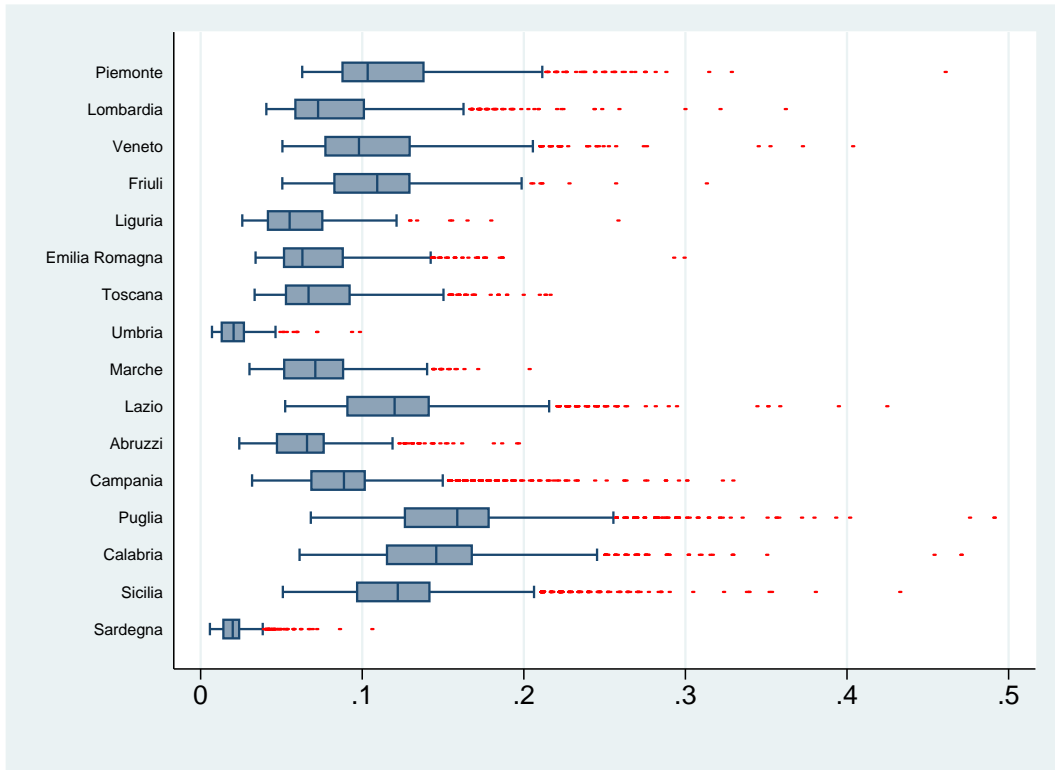
Note. Bootstrap standard errors in parenthesis. * means significant at 5 percent level, ** significant at 1 percent level, *** significant at 0.1 percent level. The regression also included a full set of provincial and year dummies. Income is in 10,000 Euros.

Table 8: Parametric estimates of the bounds on the probability of not-defaulting, South.

	Lower Bound	Upper Bound	Tightened Upper Bound
Home-owner	0.103 (0.015)***	-0.004 (0.016)***	0.030 (0.017)***
Job tenure	0.277 (0.029)***	0.035 (0.022)***	0.127 (0.023)***
Squared job tenure	-0.056 (0.010)***	-0.002 (0.007)***	-0.023 (0.007)***
Income	0.126 (0.023)***	0.013 (0.028)***	0.046 (0.016)***
Income-squared	-0.002 (0.004)	-0.000 (0.006)	-0.001 (0.002)
Length of trials	-0.081 (0.011)***	-0.081 (0.011)***	-0.098 (0.012)***
Number of branches	0.000 (0.001)	-0.002 (0.001)	-0.002 (0.001)
Self-employed	-0.184 (0.023)***	-0.162 (0.022)***	-0.186 (0.024)***
No bank account	-0.451 (0.024)***	-0.492 (0.021)***	-0.521 (0.022)***
Reliance on friends financial help	-1.825 (0.304)***	-0.214 (0.311)***	-0.851 (0.333)***
Observations	38,427	38,427	33,600

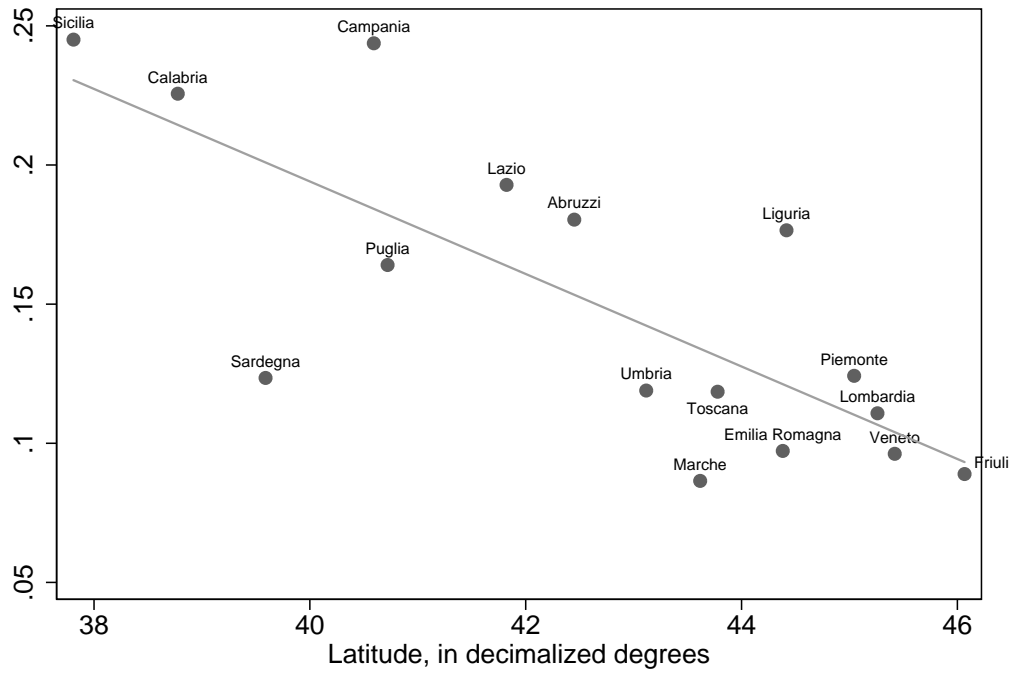
Note. Bootstrap standard errors in parenthesis. * means significant at 5 percent level, ** significant at 1 percent level, *** significant at 0.1 percent level. The regression also included a full set of provincial and year dummies. Income is in 10,000 Euros.

Figure 2: Reliance on Family and Friends



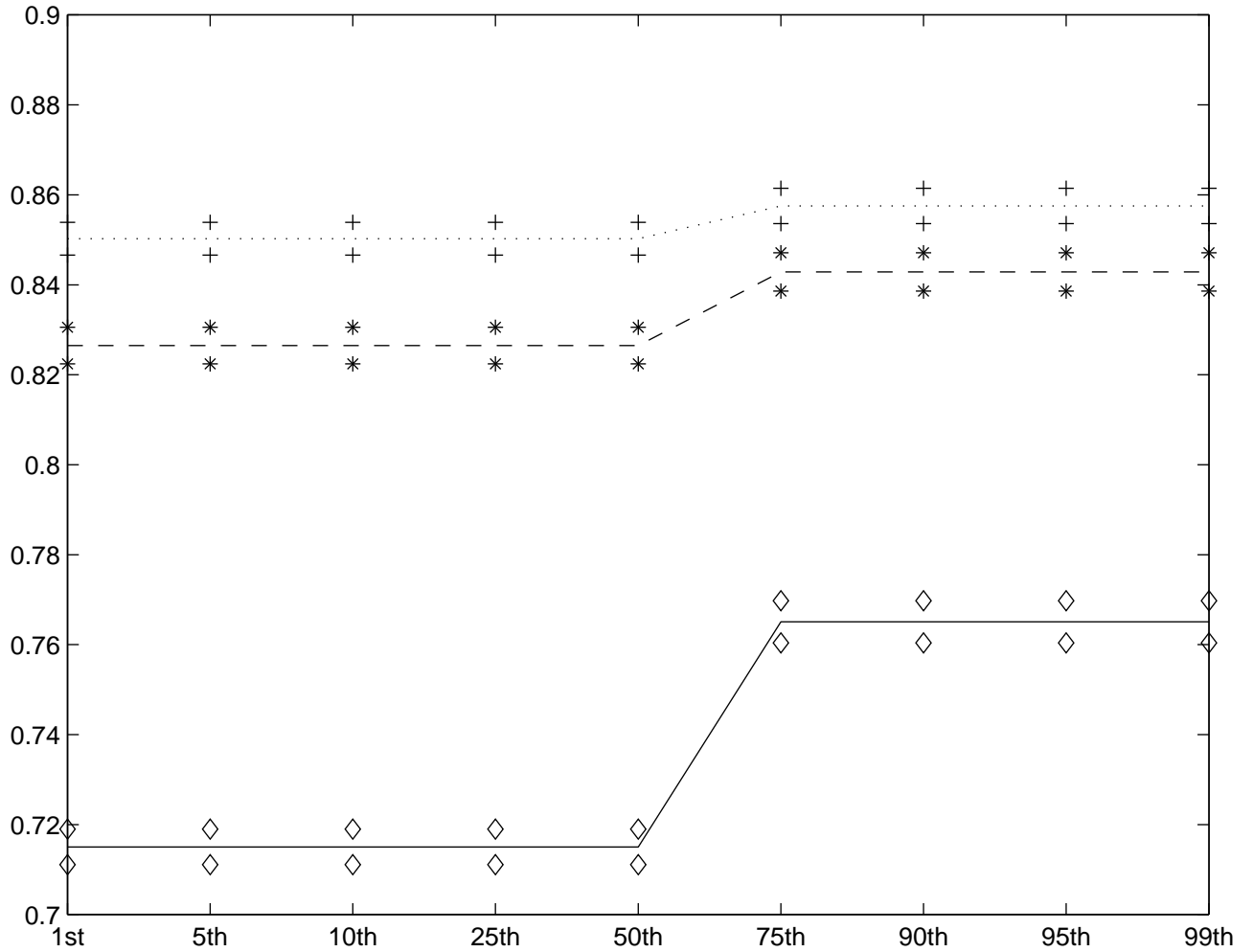
Note. The figure plots the distribution of the percentage of households reporting that they are borrowing from informal sources of credit in each Italian region. The boxes are centered on the regional median, the left hinge refers to the 1st quartile, the right hinge to the 3rd, the horizontal lines connect the 1st quartile with the lower adjacent value and the 3rd with the upper. Dots refer to observations outside the range between the lower value adjacent to the 1st quartile, and the upper value adjacent to the 3rd quartile.

Figure 3: Default and Latitude



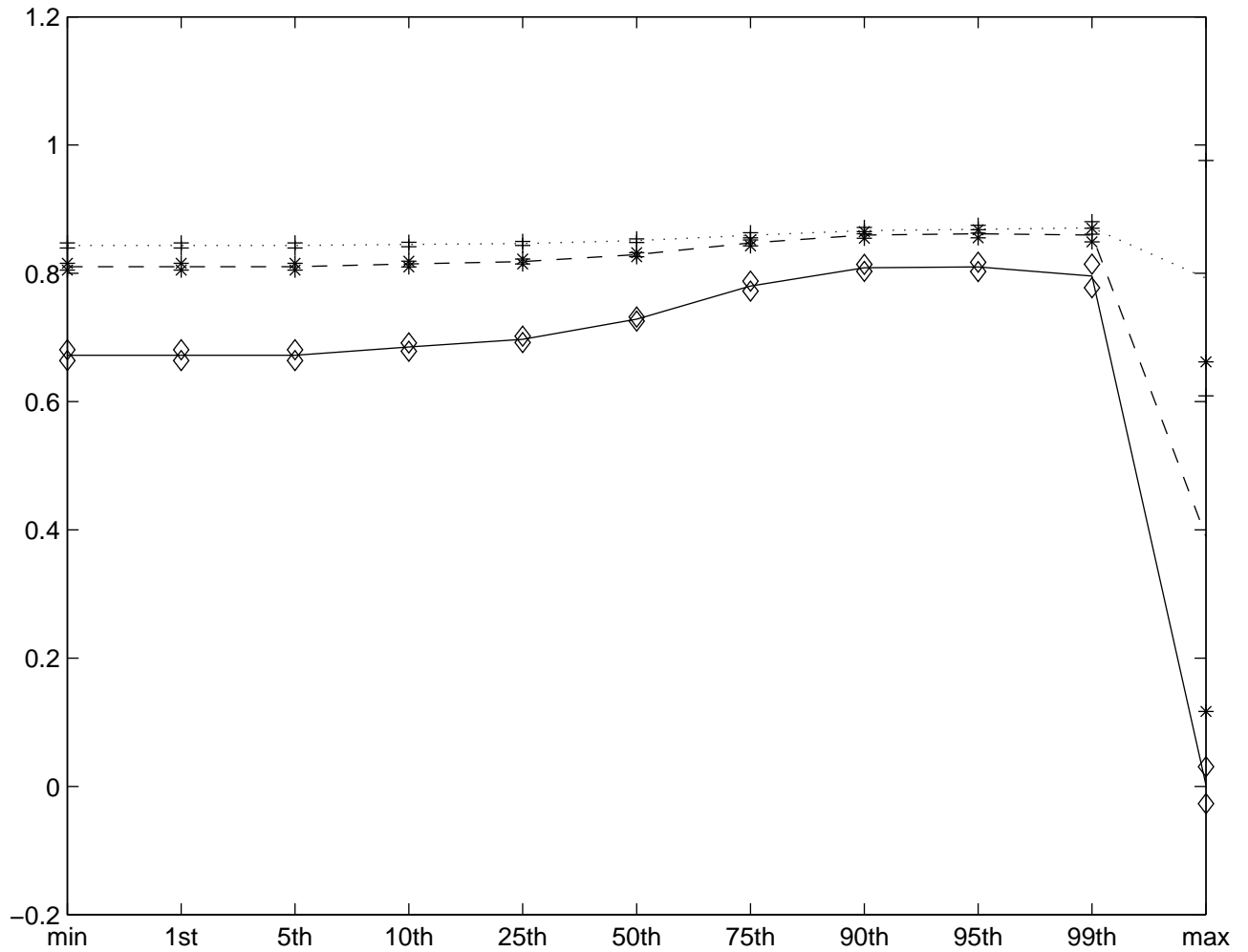
Note. The default rate is calculated using the Findomestic data and reports the percentage of contracts which have missed three or more payments when they have been due. Latitude refers to the latitude of the main city in each Italian region.

Figure 4: The probability of repaying and home-ownership: parametric estimates



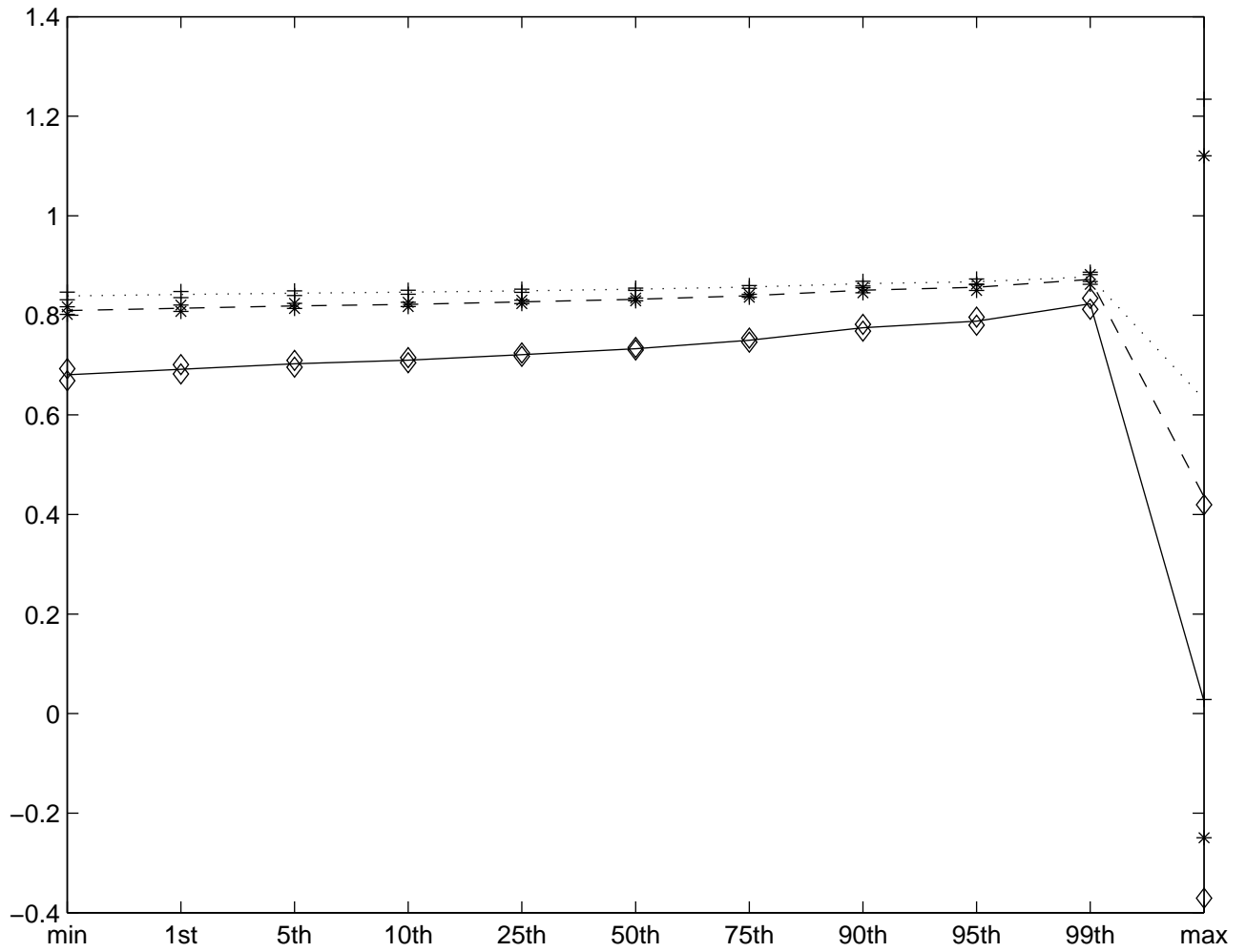
Note. Solid line represents the ‘upper’ bound, dotted line the ‘lower’ bound and the dashed line the ‘tightened upper’ bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 5: The probability of repaying and job tenure: parametric estimates



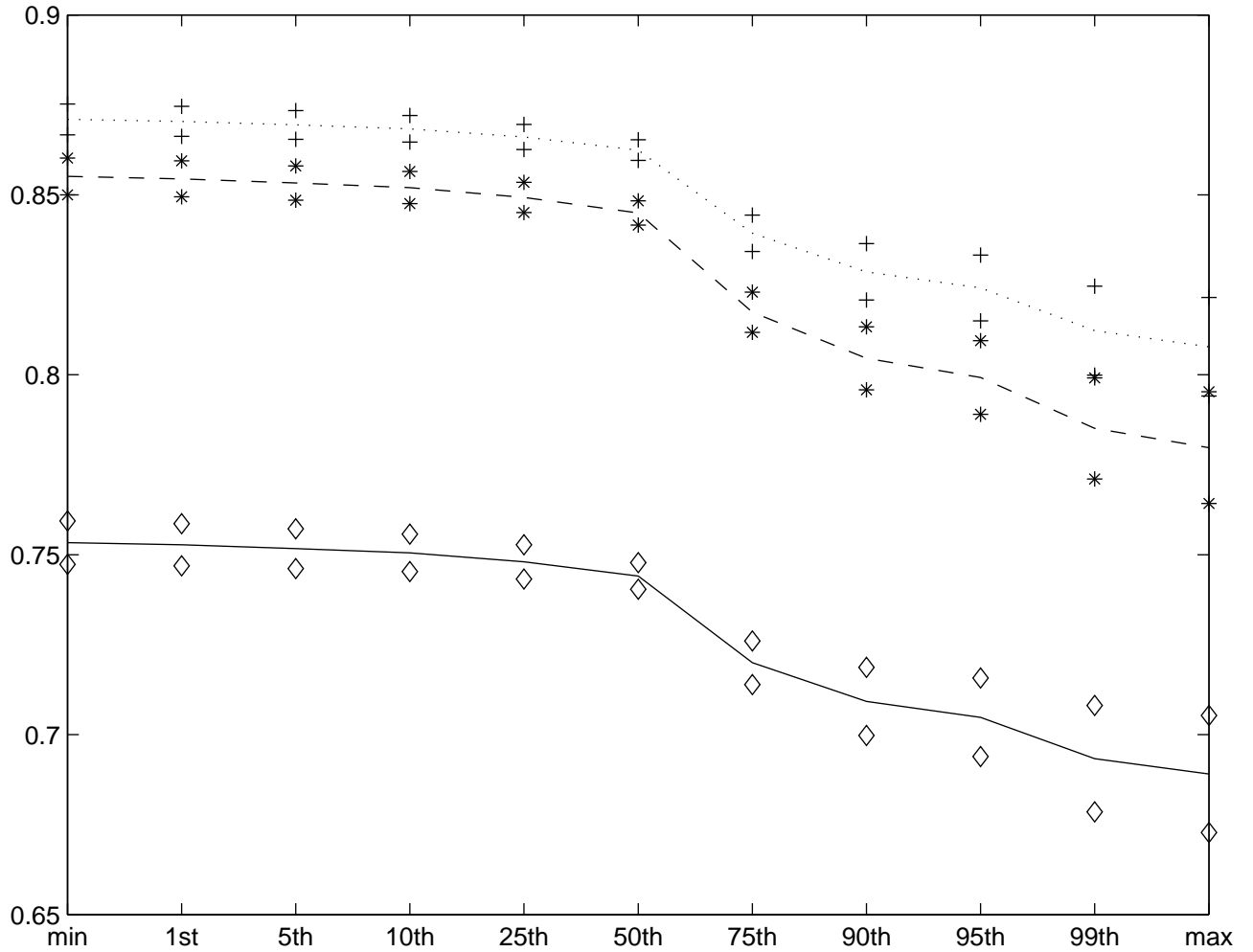
Note. Solid line represents the ‘upper’ bound, dotted line the ‘lower’ bound and the dashed line the ‘tightened upper’ bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 6: The probability of repaying and income: parametric estimates



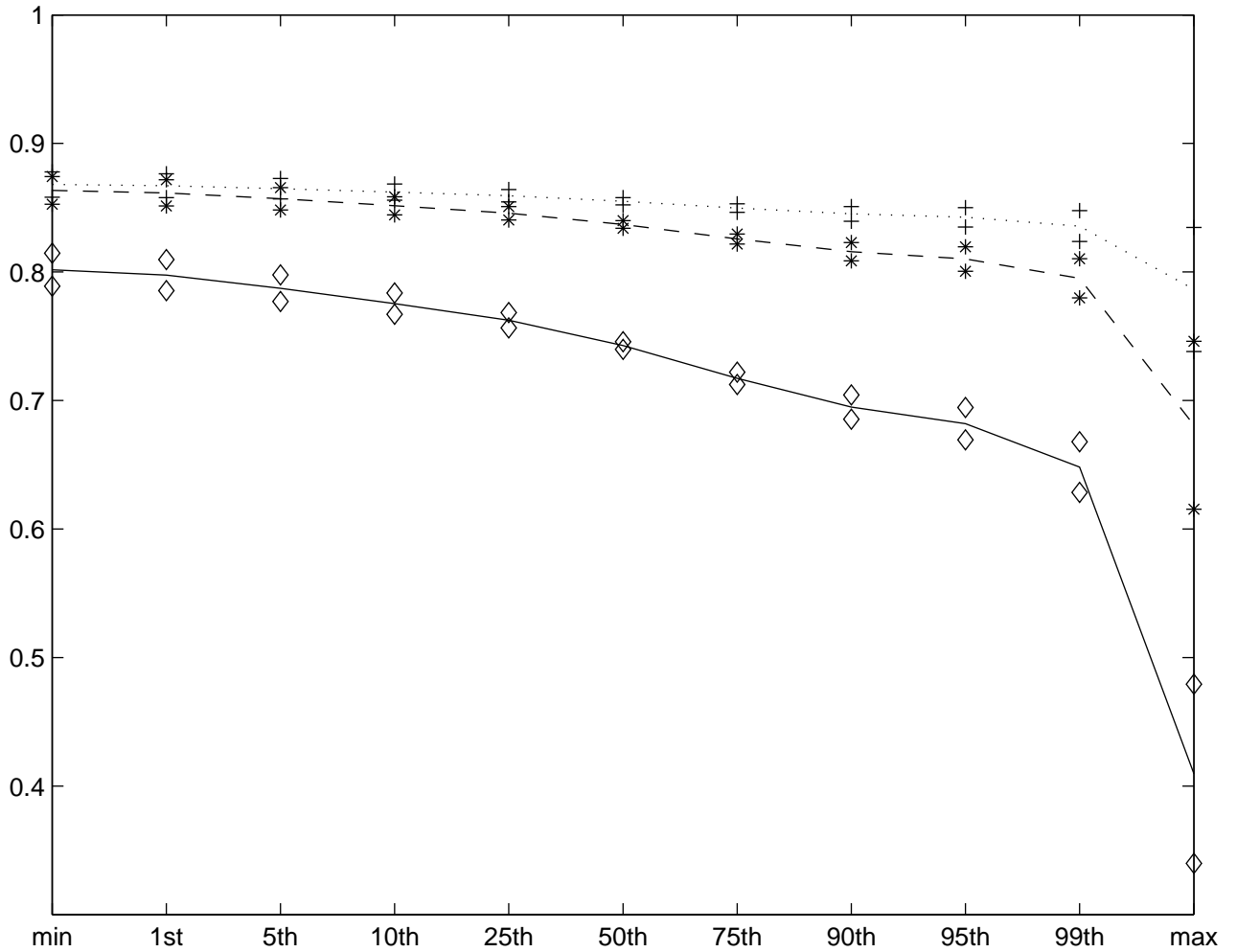
Note. Solid line represents the 'upper' bound, dotted line the 'lower' bound and the dashed line the 'tightened upper' bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 7: The probability of repaying and judicial enforcement: parametric estimates



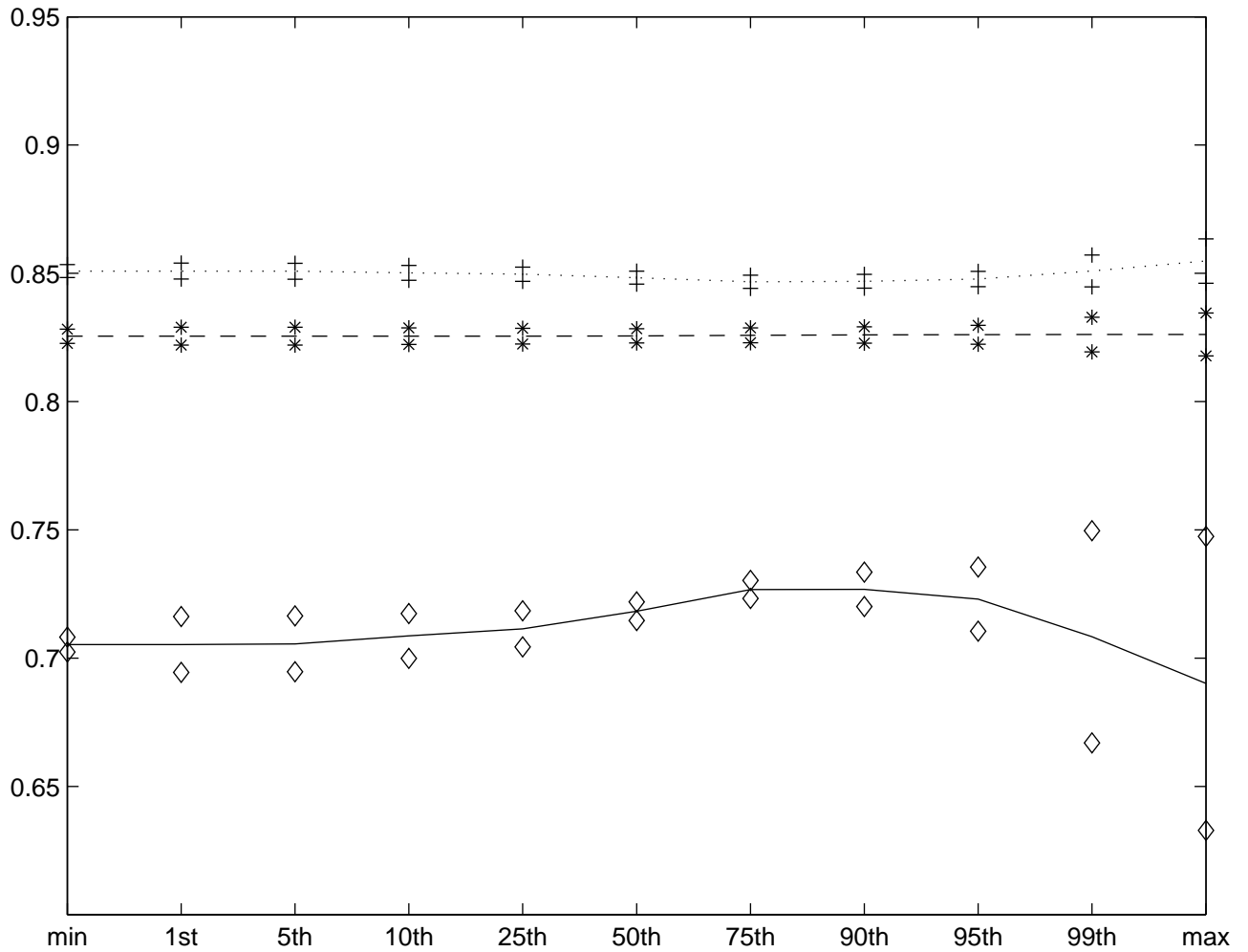
Note. Solid line represents the ‘upper’ bound, dotted line the ‘lower’ bound and the dashed line the ‘tightened upper’ bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 8: The probability of repaying and family financial help: parametric estimates



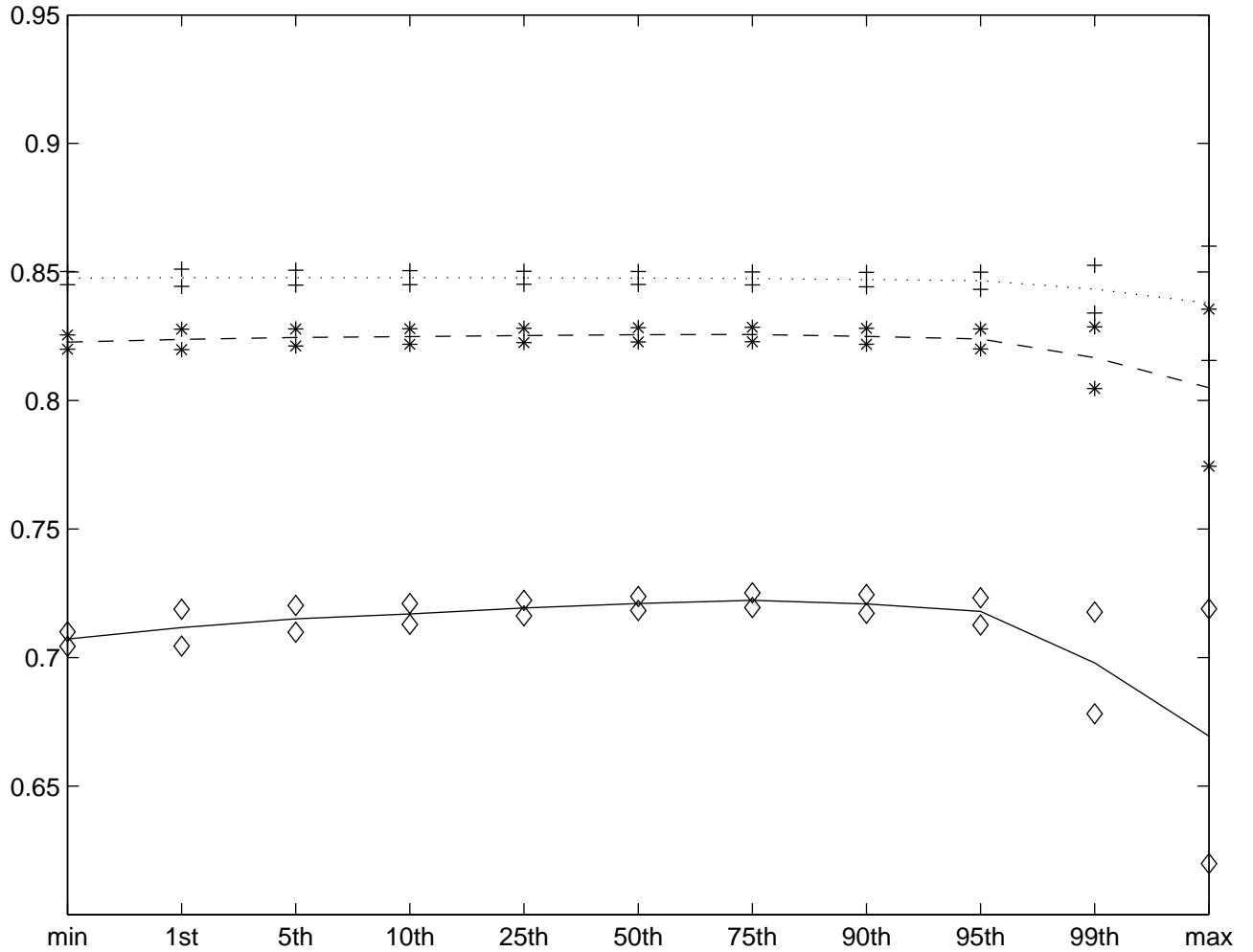
Note. Solid line represents the ‘upper’ bound, dotted line the ‘lower’ bound and the dashed line the ‘tightened upper’ bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 9: The probability of repaying and job tenure: semi-parametric estimates



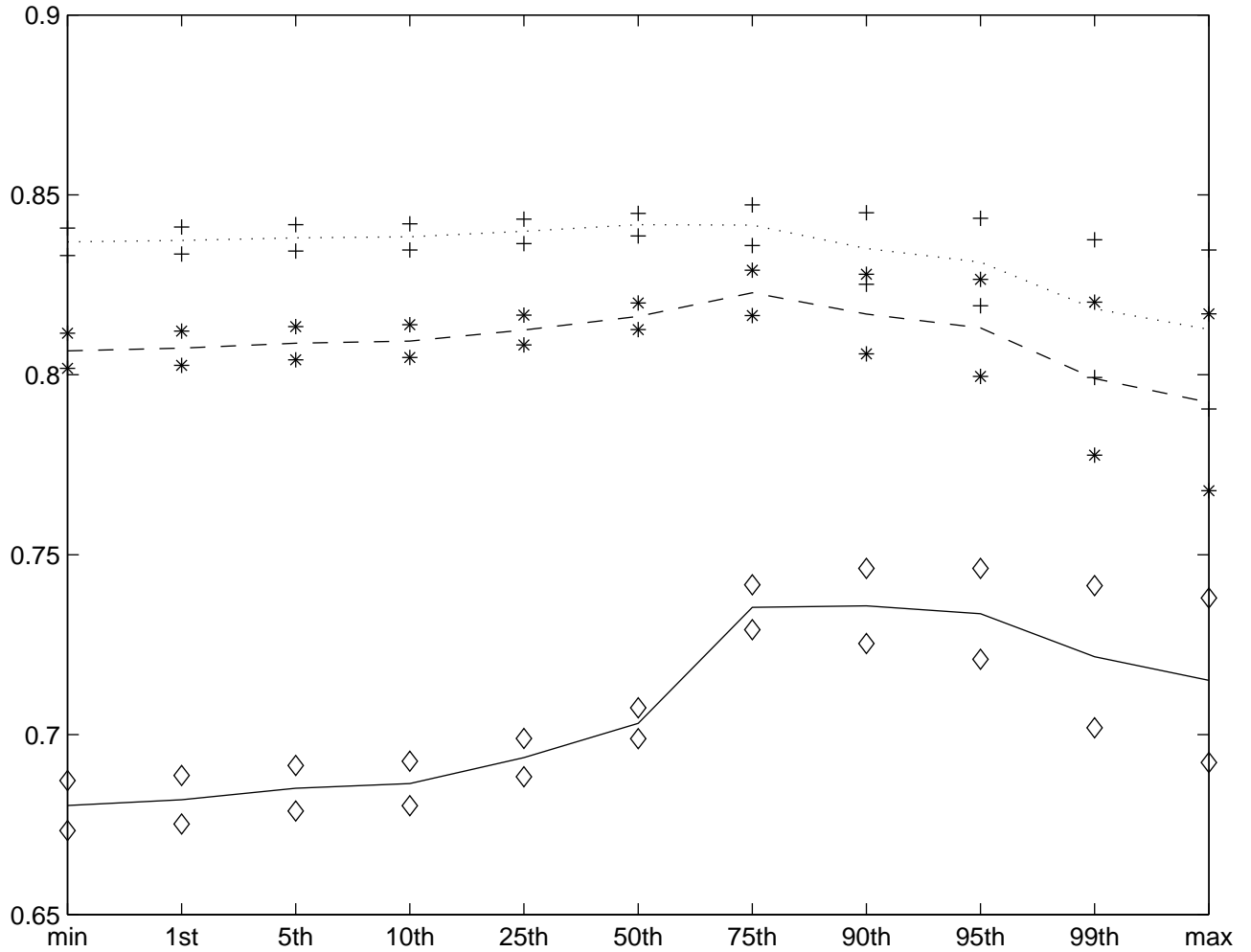
Note. Solid line represents the ‘upper’ bound, dotted line the ‘lower’ bound and the dashed line the ‘tightened upper’ bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 10: The probability of repaying and income: semi-parametric estimates



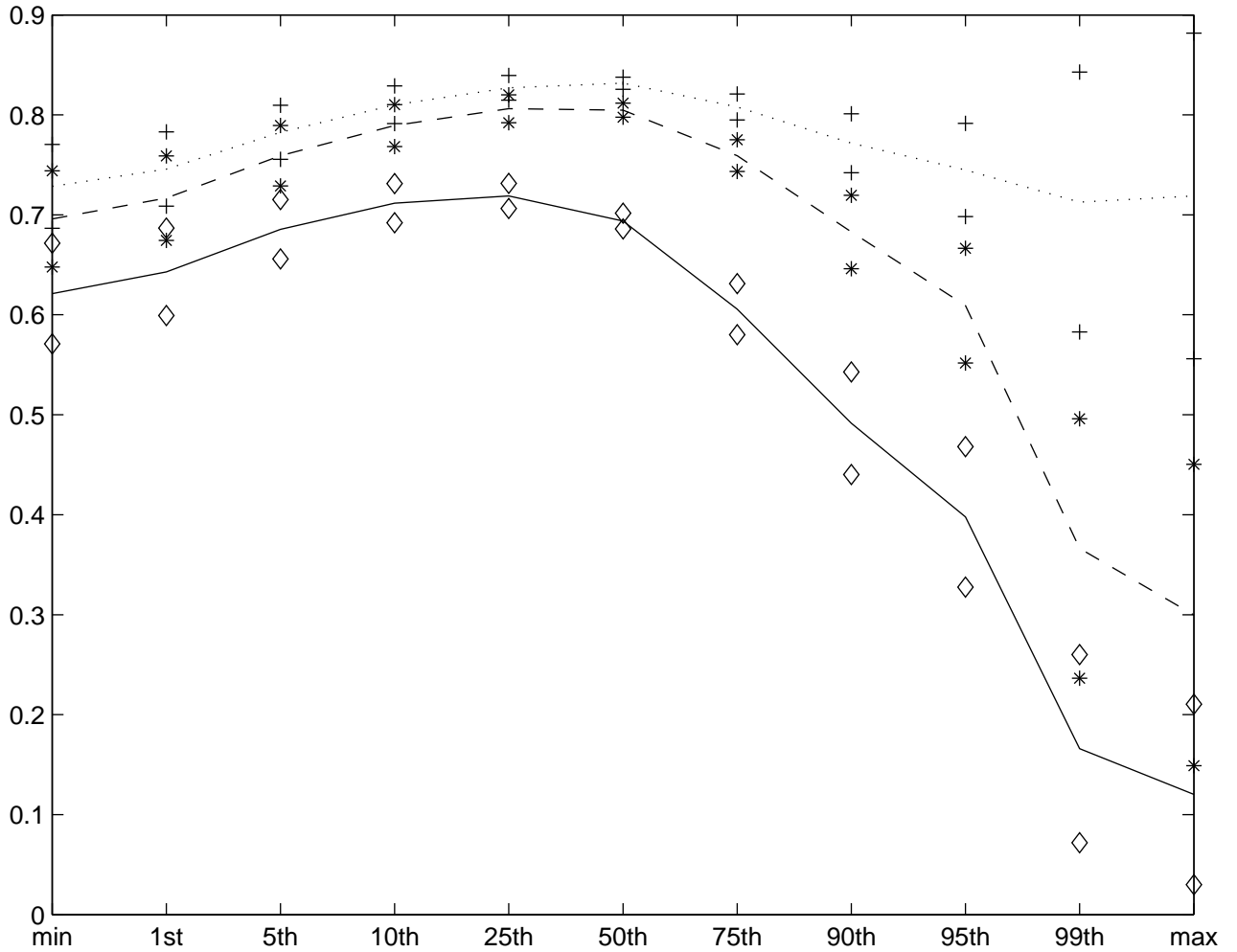
Note. Solid line represents the 'upper' bound, dotted line the 'lower' bound and the dashed line the 'tightened upper' bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 11: The probability of repaying and judicial enforcement: semi-parametric estimates



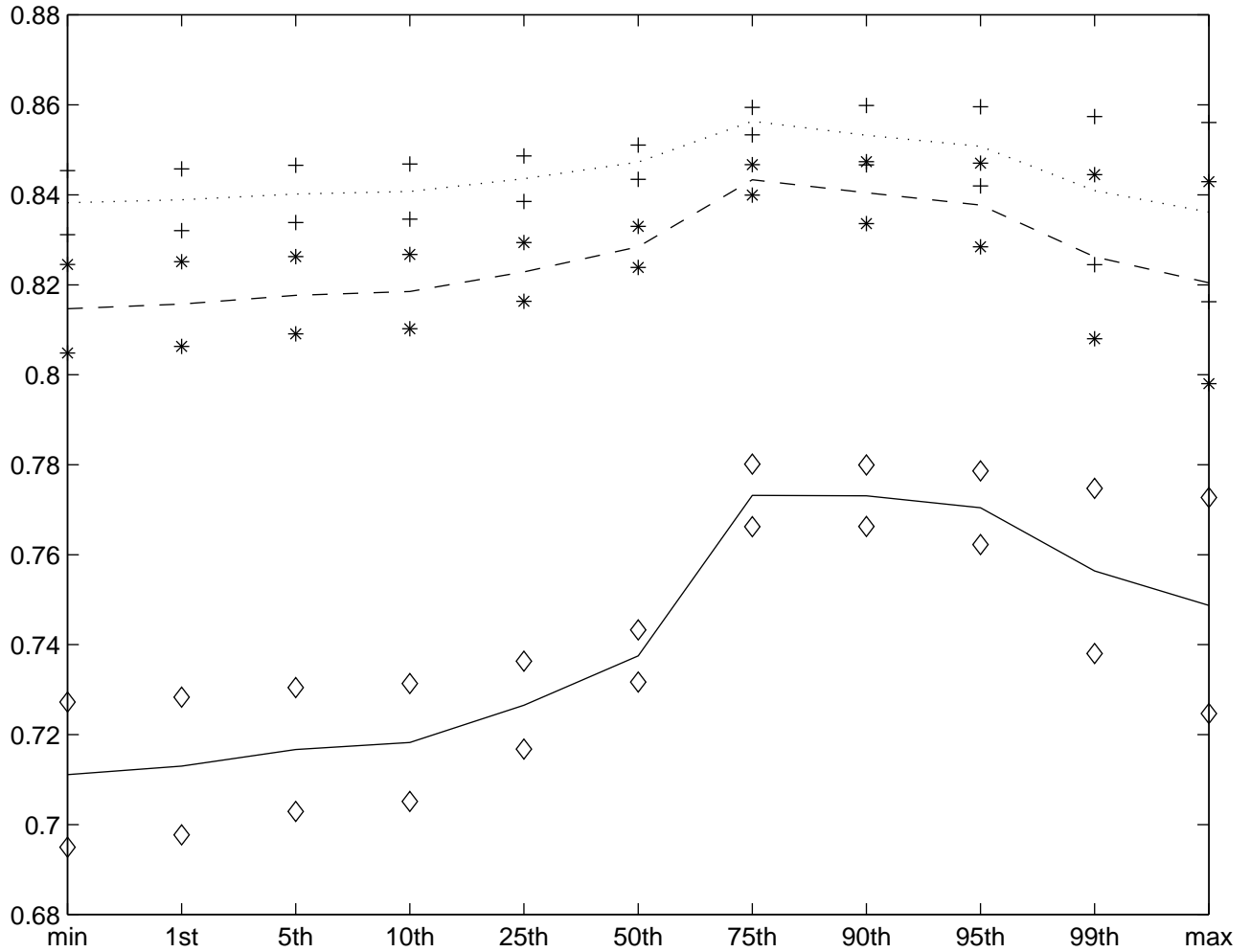
Note. Solid line represents the 'upper' bound, dotted line the 'lower' bound and the dashed line the 'tightened upper' bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 12: The probability of repaying and family financial help: semi-parametric estimates



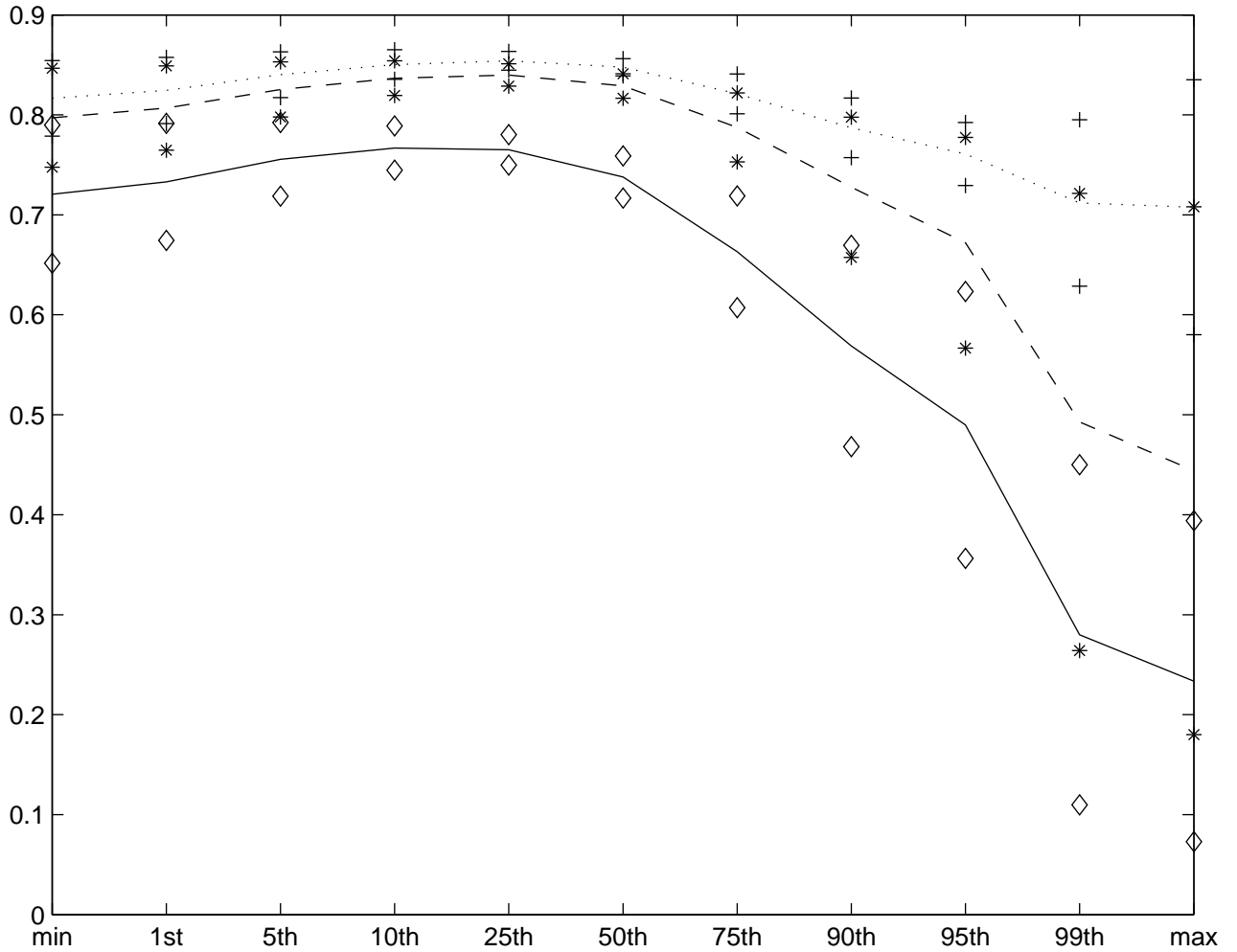
Note. Solid line represents the ‘upper’ bound, dotted line the ‘lower’ bound and the dashed line the ‘tightened upper’ bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 13: The probability of repaying and judicial enforcement: semi-parametric estimates, expired and not-expired contracts



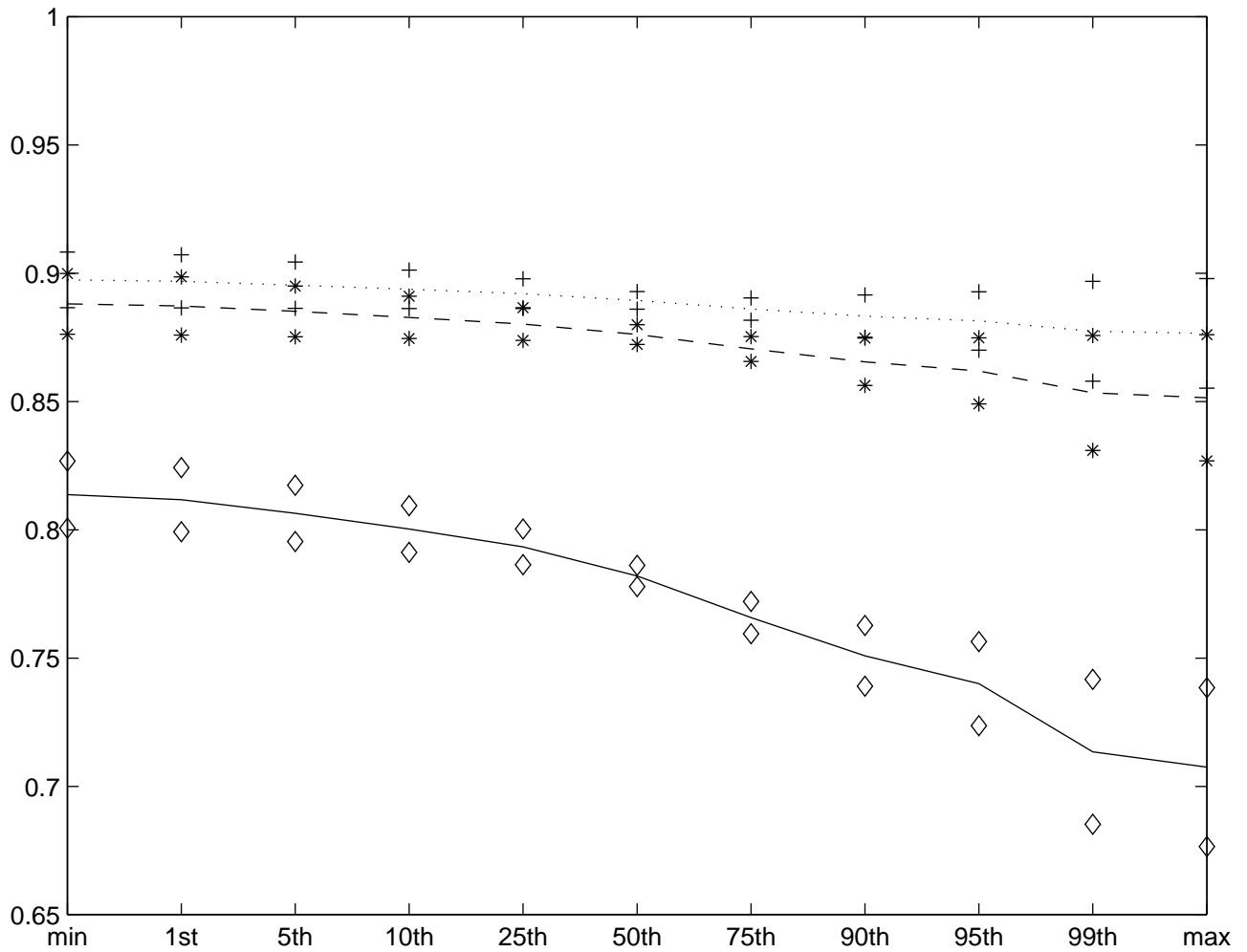
Note. Solid line represents the 'upper' bound, dotted line the 'lower' bound and the dashed line the 'tightened upper' bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 14: The probability of repaying and family financial help: semi-parametric estimates, expired and not-expired contracts



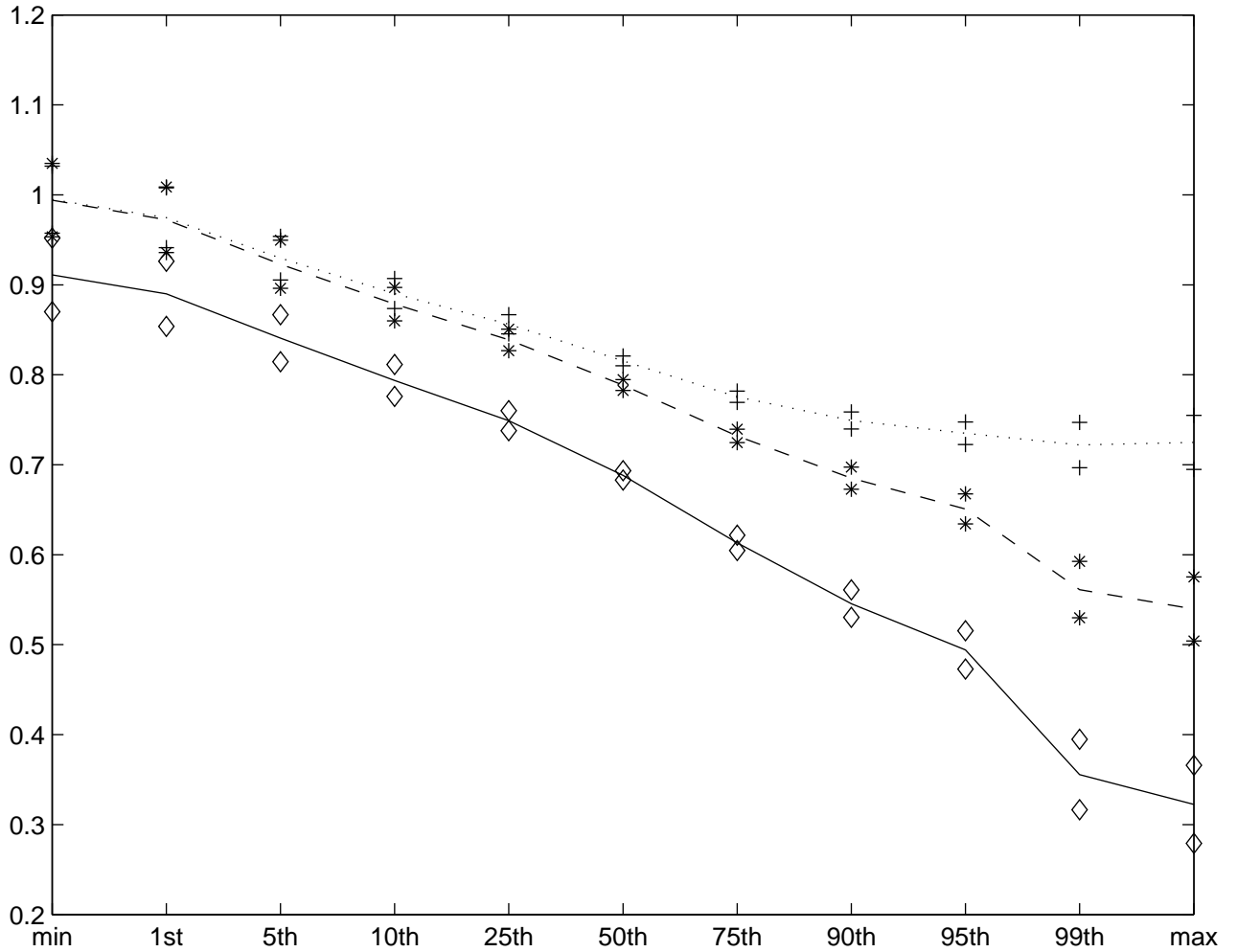
Note. Solid line represents the 'upper' bound, dotted line the 'lower' bound and the dashed line the 'tightened upper' bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 15: The probability of repaying and family financial help: semi-parametric estimates, North



Note. Solid line represents the ‘upper’ bound, dotted line the ‘lower’ bound and the dashed line the ‘tightened upper’ bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.

Figure 16: The probability of repaying and family financial help: semi-parametric estimates, South



Note. Solid line represents the ‘upper’ bound, dotted line the ‘lower’ bound and the dashed line the ‘tightened upper’ bound, with diamonds, stars, and crosses the 2 standard deviation intervals around the estimate.