Estimating Credit Constraints among US households.

Charles Grant*
European University Institute

December 2003

Abstract

Households are constrained if they want to borrow, but banks restrict their lending. This paper separately identifies (using appropriate exclusion restrictions) the demand for debt, and the maximum amount agents can borrow when it is unknown which consumers are constrained. Using data from the CEX, it estimates that between 26 percent and 31 percent of households are constrained: and that poorly educated, ethnic minority, low income, men, and (among for the educated) older households are less often constrained. On average, households would like to borrow up to $4,000 dollars more. But it does not test whether constraints are never binding.

Key Words: Credit Constraints, Consumers, Debt
(JEL C51, D12)

*I would like to thank Erich Battistin, Giuseppe Bertola, Vassilis Hajivassiliou, Maite Martinez-Granado, Nick Souleles, Mario Padula and Frank Vella whose help, support and advice enabled me to write this paper. All remaining errors, are, of course my own. Address for correspondence: charles.grant@iue.it

It is widely believed by economists that at least some agents are credit-constrained. It is offered as one of the most important explanations for a wide-variety of phenomena that are observed in economics, and also implicitly informs the policy debate, not only at the macro-economic level in motivating fiscal and monetary policy, but it also motivates micro policies such as subsidising university education for under-graduates, or supporting small business investment. While credit constraints are widely seen as pervasive, little is known about its incidence or importance. For example, while few black households borrow (as will be documented here), there is considerable debate about whether this is due to their low demand (as claimed by lenders), or to these households more often being denied credit. This paper will study credit constraints and particularly investigates consumer behaviour. It will demonstrate how to estimate credit constraints, even when all that is observed is the amount agents borrow.

When considering how much consumers borrow, several questions may be of interest. (i) Are there restrictions on the amount that consumers can borrow, that is, are at least some consumers credit constrained? (ii) How many consumers are credit constrained? (iii) How do credit constraints differ with household characteristics? (iv) How much more would these consumers borrow if unconstrained?

The approach taken in this paper is to briefly characterize what it means for a consumer to be credit-constrained, and then to use a very simple form of credit-constraints that the literature motivates to answer some of the questions that are raised above. As will be seen, a simple way to characterize debt holdings is to think of actual debt as the minimum of the amount of debt that the household wishes to hold (the demand for debt), and the maximum amount that any lender is prepared to lend (the supply of debt). While in this framework, for reasons to be discussed below, it is not possible to test the first of the questions that we might want to answer, the other questions can be answered.

The paper starts with a brief review of the literature in section 1. This discussion is restricted to the literature on consumption, and will help to motivate the rest of the paper. Section 2 proposes an estimation strategy, and explains that identification requires exclusion restrictions on the parameters that enter the demand and supply equations. Estimation must cope with two problems. First, whether a household is constrained is not directly observed, but must be replaced with some proxy variable that will be a function of observed characteristics. Second, there is a selectivity issue since even if it was observed which households were credit-constrained, demand (or supply) conditional on the household being unconstrained does not equal the unconditional demand for debt. Having estimated these equations, section 3 will then recover the estimated incidence of credit-constraints among US households, and how this differs with household characteristics. This section will also discuss how to recover a measure of how much more credit constrained households would like to borrow. The selection issue involved in estimation will also be important here as it requires the construction of the difference between demand and supply, conditional on being credit-constrained and the former can not be replaced by, for instance, the unconditional demand for debt, without downward biasing the results.

The data used is a sample of over 7,000 households from the US consumer expenditure survey for the years 1988-1993. The results show that between 26 percent and 31 percent of households are credit-constrained. This figure is higher than several previous studies. It also shows that credit constraints are more serious for single females, for well educated households and for middle income households and for white households. These differences across households are mainly due to demand, rather than supply. Additionally, credit-constraints decline with age for college educated households but are flat for poorly educated households. State banking regulations that are designed to encourage competition have little effect. Constrained households would like to borrow as much as $4,000 more.

1 The Literature On Credit Constraints.

Ever since the seminal paper by Hall (1978) consumption economists have questioned simple versions of the life-cycle/permanent income hypothesis (PIH). That paper, and an enormous number of succeeding papers, have rejected the Euler equation formulation for consumption. Another strand of the literature showed how consumption tracked income over the life-cycle, which again rejects simple versions of the PIH, for instance Carroll and Summers (1991). A number of explanations for this have been suggested in the literature: one of the most popular is that at least some consumers face binding credit constraints. These consumers would like to borrow more in order to increase their level of consumption, such consumption is compatible with their life-cycle budget constraint, but for some reason they are not able to borrow as much as they would like at the 'market clearing' interest rate. For instance, Hayashi (1987) defined consumers as credit constraints if either (i) "they face some quantity constraint on the amount of borrowing", or (ii) "the loan rate available to them is higher than the rate at which they could lend". The first is often called *credit rationing*, and there is a large literature, going back to Jaffee and Russell (1976), or Stiglitz and Weiss (1981), which motivates such credit rationing by lenders as due to the fact that some consumers default on their loans, and there is imperfect information as to which agents will default. The decision to default is not modeled. Nevertheless, such models show that it can be optimal to restrict lending to consumers. Such models imply (see figure 1) that lending takes place in discrete jumps: there are a countable number of ordered points $(0, b_1, b_2, \ldots)$ between which lending takes place at a constant marginal rate of interest. At each of these points b_k, there is a jump in the marginal rate of interest charged (perhaps to infinity, in which case no lending occurs beyond b_k).

A more recent literature has attempted to explicitly model the decision to default by consumers. This literature aimed to explain the limited ability of consumers to pool risk, and includes papers by Kehoe and Levine (1993) and Kocherlakota (1996). In these papers, the standard model of an infinitely lived, utility maximizing consumer subject to a life-cycle budget constraint is augmented by an additional constraint on the consumer's behaviour. This additional constraint explicitly accounts for the fact that *ex post* the consumer may wish to default on his debt, and suffer any penalty which ensues. The punishment for default could take many forms but these papers concentrate on default resulting in autarky, in which the consumer is permanently excluded from both borrowing and lending. By solving these models for a decentralised market economy these models can endogenously create credit rationing in which the ability to borrow is restricted to some maximum level which depends on the parameters of the model. Above this maximum level default is assured, and hence it is never rational for lenders to allow borrowing beyond this limit. Such models differ from the earlier literature in that information is perfect, and in that there is only one interest rate at which lending occurs (if it occurs at all).[1]

Several papers have estimated the proportion of households that are credit constrained. The simplest approach is taken by Hall and Mishkin (1982), who attributed the rejection of the PIH model in Euler equations to a fixed fraction of the population simply spending their current income: this fraction was estimated to be some 20 percent of the population. Hajivassiliou and Ioannides (2001) formalize how the Euler equation is affected by credit constraints, motivating their switching regression approach. Mariger (1987) tried to estimate the effective time horizon in the Euler equation and concluded that around 19 percent of households were constrained.

One problem with this approach is that the rejection of simple versions of the PIH in Euler equations could instead be due to mis-specification of the Euler equation, a point that is well known in the literature. Hence some papers have instead tried more directly to estimate or test for credit constraints. The problem is that, denoting π_i as a binary variable taking the value zero if household i is unconstrained and one if constrained, this variable π_i is not directly observed. In much of the literature, some proxy variable has been substituted for the unobserved latent variable. For instance, in a classic paper, Zeldes (1989) splits households by their level of assets: low asset households (with a gross assets to monthly income ratio of less than 2) are assumed to be credit constrained. He then documents how low and high asset households' behaviour differ. Jappelli (1990) instead uses self-reported responses to a question about credit constraints contained in the Survey of Consumer Finances.[2] The question asked if the householder had been rejected for a loan, or if he had failed to apply for a loan because he feared rejection. In either case, having chosen the proxy variable, the observations can be partitioned, and those who are thought to be credit constrained can be compared to those who are not. When a suitable proxy variable exists there is no need to estimate the incidence of credit constraints, but different groups

[1] Two papers that test the effect of punishing default are Gropp, Schulz, and White (1997) and Grant (2000). They exploited differences between state bankruptcy rules in the different US states. Grant (2000) showed that moving to a state with a stricter punishment for default results in debtors borrowing $400 more on average.

[2] Jappelli, Pishke and Souleles (1998) combines the Euler equation approach taken by Zeldes and Jappelli's more reliable measure of being constrained.

can still be usefully compared. Jappelli (1990) found that about 12 percent of households are credit constrained, rising to 19 percent if discouraged borrowers are included. He also found that credit constraints are more often binding for low income, low asset, young, and black households.

In many cases it is not clear what variable would be an appropriate proxy for credit constraints, a point made by Garcia, Lusardi, and Ng (1997). They use a switching regressions technique in the Euler equation, and note that agents should react differently to increases and decreases in income if they are at the margin of being constrained. Their technique allows for constraints to be a function of several variables, and they find that around 16 percent of agents are constrained.[3] Gross and Souleles (2002) look at credit card balances and limits, and note that consumers increased their borrowing in response to any increase in their credit limit: they interpret this as due to credit constraints and argue that the effect of credit constraints on consumer behaviour is substantial.

Early estimates of the extra amount that households wanted to borrow, such as Hayashi (1985) and Mariger (1987) found, by using an Euler equation approach, that credit constraints had little effect on debts holdings and consumption. Cox and Jappelli (1993) study a cross section of households and compare a group who are assumed to be credit-constrained (based on responses to a question about having been turned down for credit), with a group who are not. They find that constrained households would like to hold over $8,000 more debt than they actually do. Two other papers worth mentioning are Perrudin and Sorensen (2000) and Duca and Rosenthal (1994). The first considers a two stage estimation of asset holding using the 1983 wave of the SCF, whereby a probit predicts which asset types are held, while the second stage predicts how much of each asset is held. They find that age, marital status, education, and sex all have substantial effects on both the type and quantity of assets. In contrast, Duca and Rosenthal (1994) look at how liquidity constraints affect the ability of households to enter the mortgage market again using Jappelli's 'turn-down' measure. Their model allows for selectivity by using a bivariate probit model for the housing choice and whether a household is credit constrained, finding borrowing constraints

[3] As in this paper, their approach imposes that the probability of being credit constrained is bounded away from zero for the regression to be identified.

particularly affect younger households. They ask whether households borrow, but not how much.

One problem is that even if it were known which households were credit constrained, estimates of the demand, or other behavioural equations can still be biased. Estimates of the demand equation based only those observations who are unconstrained (for which $\pi_i = 0$) are likely to under-estimate the true demand for debt among constrained households. Those households with an unusually low level of demand, in the sense that they have low error draws in the demand equation, are less likely to be observed to be credit constrained. This selection problem must be accounted for when recovering a true estimate of how much more credit-constrained households want to borrow than they are currently allowed. This problem occurs when Cox and Jappelli (1993) estimate how much more households wish to borrow, but is controlled for by Duca and Rosenthal (1994).

The next section is devoted to showing how to solve the identification and selection problem when no good proxy variable is available. It directly models credit constraints in an econometric model, and discusses how this model can be identified. In particular, credit constraints are not replaced by some proxy variable but instead they are modeled as arising from some equation $\pi_i = f(X_i, \varepsilon_i)$ where X_i is continuous and multi-dimensional. Furthermore, it is implicitly recognized that π_i is observed with error.

2 An Empirical Framework

Theory suggests (recall figure 1) that consumers can borrow any amount up until some limit at which the household is constrained. Both the amount that the consumer wishes to borrow, and the credit limit are functions of the household's characteristics denoted X_i, where i denotes the household. Describing desired borrowing, y_{1i}, as "demand" and the credit limit, y_{2i}, as "supply" then the supply and the demand for credit can be written as functions of the households characteristics:

$$\begin{aligned} y_{1i} &= f_1(X_{1i}, \varepsilon_{1i}) \qquad && \textit{demand} \\ y_{2i} &= f_2(X_{2i}, \varepsilon_{2i}) \qquad && \textit{supply} \end{aligned} \tag{1}$$

where ε_{1i} and ε_{1i} are the unexplained errors, or stochastic parts of the demand and supply functions respectively. The actual level of debt observed, denoted y_i, is defined as the minimum of supply and demand (or zero if this number is negative). These equations explicitly recognize that the econometrician does not observe all the characteristics that drive demand or supply. For an agent to be credit constrained both demand must exceed supply, and demand must be positive. This formulation makes explicit that credit constraints can bind at some level other than zero, and that not all agents who do not borrow fail to borrow because they are credit constrained. An alternative and parsimonious way of representing the same result (ignoring the zero observations) is to write

$$y_i = f_{1i} + \pi_i \left(f_{2i} - f_{1i} \right) + \varepsilon_i \tag{2}$$

How estimation proceeds depends on what exactly is observed. The approach taken by Cox and Jappelli (1993) assumes that π_i is observed, or can be well approximated by some proxy variable which partitions the data into those for whom supply is binding and those for whom it is not. From those who are credit constrained (where $\pi_i = 1$) the supply equation can be recovered, while the demand equation can be recovered from those who are not. In this case the extra amount that credit-constrained consumers wanted to borrow is the difference between demand and supply conditional on credit constraints being binding. The strategy highlighted replaces demand conditional on credit constraints by the estimated demands from those for whom the constraints is not binding. However, estimation must account for the selection problem: agents are only observed to be credit constrained (ignoring the zero observations for the time being) if demand exceeds supply. That is, any estimation strategy must explicitly recognize that

$$E \left(\varepsilon_{1i} | \pi_i = 0 \right) \neq E \left(\varepsilon_{1i} | \pi_i = 1 \right) \tag{3}$$

and similarly for ε_{2i}. Failure to account for this selection problem results in biased estimates of f_1 and f_2 and thus mis-estimates how much more credit-constrained consumers wish to borrow.

More often π_i is not observed, and there is no good proxy variable to replace it. In these cases the problem is that only y_i is observed, but it is not clear *a priori* which of the underlying equations y_{1i} or y_{2i} has generated the observation. However, estimation can proceed by noting that what is observed is the minimum of supply and demand. That is:

$$y_i = \begin{cases} min\,(y_{1i}, y_{2i}) & y_{1i} > 0, y_{2i} > 0 \\ 0 & otherwise \end{cases} \tag{4}$$

This involves using the estimated function $\pi_i = f\,(X_{1i}, X_{2i}, \varepsilon_i)$. Two comments are worth making. First, while previously the proxy for π_i was binary and one dimensional, this formulation explicitly recognizes that many variables can affect the incidence of credit constraints, and that such variables may be continuous. Secondly, it also recognizes that there will be heterogeneity across agents that is not captured through those variables X_{1i} and X_{2i} observed by the econometrician. This fact is captured by the addition of the error term ε_i. The aim is to replace π_i by a probability distribution which depends on the household's observable characteristics, rather than by some proxy variable. This probability depends on the estimated parameters in the underlying supply and demand equations.[4]

A: Estimation by Maximum Likelihood

The framework discussed above, for which π_i is not observed, is very similar in form to standard canonical disequilibrium models, as discussed in Quandt (1988), and the discussion is similar to that contained there. Estimation can proceed by full maximum likelihood. The likelihood of any observation y_i is thus

$$\begin{aligned} \mathcal{L}_i = \; & Pr\,(y_i | y_{1i}, y_{2i} \geq 0; y_{1i} < y_{2i})\, Pr\,(y_{1i} < y_{2i} | y_{1i}, y_{2i} \geq 0)\, Pr\,(y_{1i}, y_{2i} \geq 0) \\ & + Pr\,(y_i | y_{1i}, y_{2i} \geq 0; y_{1i} \geq y_{2i})\, Pr\,(y_{1i} \geq y_{2i} | y_{1i}, y_{2i} \geq 0)\, Pr\,(y_{1i}, y_{2i} \geq 0) \\ & + [1 - Pr\,(y_{1i}, y_{2i} \geq 0)] \end{aligned} \tag{5}$$

where it is implicitly recognized that we are also conditioning on X_{1i} and X_{2i}. The first part of the equation represents the contribution to the likelihood function of those observations where debt is positive, and in which supply exceeds demand so that the household is unconstrained. The second part represents those observations for which demand exceeds

[4] If repeated observations of the same household are available then a variety of different techniques are available, that can exploit the panel structure of the data. See for instance Eaton and Gersovitz (1981) or Hajivassiliou (1987).

supply, while the final part represents those households who do not borrow. If we impose that f_1 and f_2 are linear, and further assume that the error structure is bivariate normal with co-variance matrix Σ, so that

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim N \left[0, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right] \tag{6}$$

and note that y_i is observed, then it is simple to construct the likelihood function for this problem. Although the exact form of the likelihood is given in the appendix, some technical remarks need to be made. Consistency of the estimator follows from the consistency of both the tobit model and of the canonical form of the disequilibrium model: for the later Hartley and Mallela (1977) highlighted a number of conditions that are needed to ensure the consistency of the estimator. The most important of which are (i) that the parameters to be estimated are in the interior of the parameter space; and (ii) there exist exclusion restrictions on the supply and demand equations. A third condition is that the fraction of observations falling within each regime approaches a strictly positive fraction as the number of observations approaches infinity. These conditions translate directly to the estimator above. The first of these conditions means that the estimated variances σ_1 and σ_2 must be bounded away from zero, and that the estimated correlation parameter of these errors ρ must be bounded away from ± 1. This is more onerous than might be thought: unlike in the straightforward tobit model, it is not known, a priori, whether a non-zero observation obtains from the supply or the demand equation. If the regression includes a constant, the coefficients on the parameters can always be chosen so as to make $f_1 = 0$ in which case the likelihood is unbounded as $\sigma_1 \longrightarrow 0$ (similarly for supply). In practise this problem translates into finding a suitable starting value that does not lie in the region in which the gradient points to the boundary of the parameter space.[5] The second condition means that there must be variables that enter the supply, but not the demand equation, and/or variables that enter the demand, but not the supply equation (e.g. $X_{1i} \neq X_{2i}$). This is the identification problem. Note however, that given such variables, other variables can freely

[5]This was done by using the estimated parameters from a tobit model as the initial starting values. Kooiman, van Dijk and Thurik (1985) suggest instead employing some penalty function for getting too close to the boundary, although it is not clear why this should cause convergence to the interior maximum.

enter both the supply and the demand equation, and the estimated effect on supply and/or demand of the variable can be separately identified.

Which exclusion restrictions should be made? The exclusions in this paper are that quarter enters demand and not supply, while bank regulation, and number of people per bank in the state that year both enter supply and not demand.[6] Are these restrictions reasonable? Using seasonal dummies only in the demand equation argues that lenders do not discriminate on the basis of which month borrowers ask for loans.[7] If banks did, then there would be incentives for borrowers to time their requests for debt at certain times in the year, and since it is as costly to request a loan in one month as in another, it seems sensible to suppose there is a pooling equilibrium on month. Furthermore, there is some evidence that the federal reserve manages monetary policy to eliminate seasonality in supply: this is a movement along the supply curve rather than of the supply curve, see Miron (1986) or Barsky and Miron (1989). On the other side, the banking regulations are due to state banking laws.[8] These rules, tabulated in table 1, not only vary across states but also across time, and it seems reasonable to suppose that they are unrelated to demand. The table shows when different states allowed intra-state branching, and highlights that there has been a gradual relaxation in banking regulations over the last 20 years. The timing of this de-regulation differed from state to state, which can be exploited in the regressions that will be run. The regressions will also include people per bank in the state as a proxy for local bank competition.

---

[6]Strictly speaking, we only need either something in demand and not supply or something in supply and not demand, and not both. Further to the exclusion restrictions reported above, some regressions are also reported with additional exclusion assumptions.

[7]Note that this does not mean that debtors observed, or unobserved characteristics do not change month by month, merely that banks do not use month in their assessment of whether to extend a loan to the potential debtor.

[8]The rules used are a dummy for whether intrastate branching through merger and acquisition is allowed (law 1) and a dummy for whether full intrastate branching was permitted (law 2). Including whether interstate banking was also permitted made no difference to the regressions. These rules were obtained from Amel (1993) and Kroszner and Strahan (1999) where a much fuller discussion of these regulations is contained.

The third condition means that at least some, but not all, households are credit constrained in the sample. Moreover, in reported regressions assumed that the errors ε_{1i} and ε_{2i} in equation 1 are uncorrelated ($\rho = 0$), since convergence failed when this restriction was not imposed. One interpretation of the error term is that it is due to parameters being omitted by the econometrician from the regression, either because they are unmodelled, or because they are not observed. For the parameters in the regression to be identified it must be true that $E\left(\varepsilon_{ji}|X_{ji}\right) = 0$ for $j = 1, 2$, the standard assumption in a regression. However, if $\sigma_{12} = 0$ then this argues that any such omitted variable enter either the supply equation, or the demand equation, but not both. While this seems unlikely, we hope the imposition of this assumption will not affect the results too strongly. If instead an omitted variable increased (or decreased) both supply and demand, then this would cause a positive correlation between the errors, in which case $\sigma_{12} > 0$, while if the omitted variable entered the supply and demand equations with opposite signs, then $\sigma_{12} < 0$.

3 Data Description:

The data that this paper uses is the Consumer Expenditure Survey (CEX): a survey of US households conducted by the Bureau of Labor Statistics, primary for the purpose of estimating inflation rates. Households are continuously surveyed about their consumption, together with a number of household characteristics. Since 1988, households have also been asked detailed questions about their financial position, and in particular about all their outstanding unsecured debts (including credit-card debts, debts on store cards, bank debt, debts at savings and loans companies, debts at credit unions, debts with finance companies, medical debts, and debts with other sources). This paper concentrates on unsecured debts: the debt measure it uses is the sum all these unsecured debts reported in the CEX, and excludes secured borrowing on mortgages. If households are constrained, it seems much more likely that their unsecured borrowing is constrained, rather than borrowing to buy collateralized assets. While households are interviewed 5 times (the survey is conducted as a rotating panel in which one fifth of households drop out of the survey each quarter, and are replaced by a new household), questions on debt are only held in the first and last interview. Only one of these interviews will be used since otherwise there are potential correlations between observations when the same household is observed multiple times.

Other surveys of household debt exist, but there are a number of advantages in using the CEX. The first is the large sample size, which is important given the estimation strategy. A second advantage is that the CEX survey contains information on the state of residence of the household. This is crucial since the supply side instruments are the state level banking regulations and the people per bank in the state. Few other surveys supply such information. For instance, the Survey of Consumer Finances is perhaps a more natural data source for information on household debt but it only provides state information for 1983. This means that changes over time can not be captured. Furthermore, since the SCF is conducted over a month or two, month can not act as a demand instrument using that survey. The CEX survey, by contrast, surveys continuously throughout the year, and provides state information for all years that are surveyed. For the regressions the debt was deflated by a household specific Stone-Geary price index (where the prices relate to non-durable expenditure). As is common in regressions of consumption and borrowing, the aggregate interest rate is included in the regression: we use the tax free municipal bond interest rate.

The data used in this paper are those observations from 1988 to 1993 for which full state information is available and who were 'full income responders'. To ensure a reasonably homogeneously defined group, sampling was restricted to single households or those headed by a couple, but excluding other more complex types of households. The paper also only includes those households whose head was between 25 and 55. Large households (with 7 or more members) were excluded, as well as households whose head received no education. Also excluded were self-employed households, those whose primary occupation as farming, and those households in which more adults other than the household head (and his or her partner) were working.

Table 2 summarises some of the features of the raw data, without conditioning on any observable characteristics of the households. It shows that the median debt holding in the whole sample is $736 while 68.0 percent of households hold at least some debt. Conditional on holding debt, the average amount of debt held is $3,984, a substantial amount. While this may seem large, other studies, such as Cox and Jappelli (1993) have found similarly large amounts. The table also compares the level of debt for each year. Except for the first year, the proportion holding any debt gradually declined over the period in question. However, the average size of the debt was increasing, and there were particularly dramatic increases in this quantity in 1991 and 1993. Comparing age groups shows that younger households are more likely to hold debt than older households, and that median debt holdings are roughly twice as large. The level of debt was for the 25-35 age group and was over \$400 more than for the 45-55 age group but \$325 less than the middle aged group. The table also highlights some other features of the data. Childless households are less likely to hold debt, and hold less debt when they do by roughly \$500. While there is little difference between households with either one or two children, having three or more children results in around \$100 less being held, and 2-3 percent fewer households hold debt. The differences between education groups is dramatic; the most poorly educated group is 25 percent less likely to hold any debt, although when they hold debt, their holdings are similar to the middle two education groups. The most educated group hold much greater amounts of debt, over \$2,000 more than any other group. Other comparisons show that unmarried households (either headed by women or by men) and ethnic minority households are all less likely to hold debts, and hold smaller debts when they do. Low income households are also much less likely to hold debt. This may seem to suggest that those groups, such as unmarried households, black households, or low income households have more difficulty in smoothing consumption since they are less likely to use the credit markets, but of course it is difficult to dis-entangle supply and demand effects from these raw numbers. The rest of the paper is devoted to this issue.

4 Results

Table 3 displays the estimated supply and demand equations. Omitted from the table are the coefficients on the year dummies and a set of regional dummies. Results are recorded for both the levels (in regression A) and for the log-levels (in regressions B, C, and D). For the log-level regressions the left-hand side variable is the $\ln(1 + debt)$. All four reported regressions imposed that the errors from the supply and the demand equations were uncorrelated. The results show that being married, getting older, and more education significantly increase demand according to the demand equation (quarter is also significant). Interestingly, black and male households have significantly lower demand. In the supply equation only marital status and having three or more children is significant. Similar results are obtained when the supply and demand equations are estimated in log-levels, the difference is that now three or more children enter demand but not supply, while education enters supply. Age is no longer significant in either demand or supply. This result motivated regression C which not only includes a quadratic in income, but also interacted the age polynomial with education. The motivation for this is the difference in the age-income profiles for highly and poorly educated households: the steeper earnings profile for educated households may result in higher demand (and/or supply) for these households when they are young. The results show that age now is important in the demand equation only when interacted with education. By contrast, income is highly significant in both the supply and the demand equation. Lastly, in regression D, race and sex are removed from the supply equation since it is illegal to discriminate between customers on these grounds (and in any case they are never significant). Moreover, since education may in practise be difficult to observe, it too has been excluded from the supply equation. The pattern of the results in this regression is similar as in regression C: again race, education and its interaction with income, sex, marital status, and income are all significant in demand, while marital status and income are significant in supply.

The raw results are themselves not very easy to understand. However, having these results allow examination of the questions that were outlined in the introduction. The first question was to test whether $\pi_i \equiv 0$, and as discussed in the appendix, is not investigated in this paper but the other questions can be addressed by constructing the following:

(ii) $E(\pi_i)$

(iii) $E(\pi_i | X_i)$

(iv) $E(y_{1i} - y_{2i} | \pi_i = 1) \, Pr(\pi_i = 1)$

Item (ii) will give (assuming our observations are a random draw from the whole population) the fraction of households credit-constrained in the whole economy. One could also (but have not) constructed the probability of being credit-constrained conditional on the current level of debt. Ranging over all possible y_i would enable a calculation of the 'maximum' level of debt that a consumer is allowed to hold. Instead (iii) gives the probability of being credit-constrained conditional on the X-variates. By ranging over the X-variates different subgroups can be compared to see if there are significant observable differences across these subgroups in their ability to borrow and smooth consumption. Lastly (iv) shows how much more such consumers would have borrowed in the absence of binding credit-constraints.

A: *The proportion of households that are credit constrained.*

The proportion of households that are credit constrained is the unconditional expectation of π_i over all households. For a household to be credit constrained the demand for debt must exceed supply *and* demand must be positive. These two conditions will not be independent, even when the supply and demand equations have uncorrelated errors. This condition can be written as

$$\begin{aligned} E(\pi_i) &= Pr(y_{1i} \geq 0; y_{1i} \geq y_{2i}) \\ &= Pr(y_{1i} \geq y_{2i}; y_{1i}, y_{2i} \geq 0) + Pr(y_{1i} \geq 0; y_{2i} \leq 0) \\ &= Pr(y_{1i} \geq y_{2i} | y_{1i}, y_{2i} \geq 0)\ Pr(y_{1i}, y_{2i} > 0) \\ &\qquad + [1 - Pr(y_{1i}, y_{2i} > 0) - Pr(y_{1i} < 0)] \end{aligned} \tag{7}$$

Not all households that have no debt wish to hold debt, and at least some of the households who hold debt would like to hold more. This probability can easily be constructed and more details are given in the appendix. The variance can also be calculated, since the estimated $\hat{\pi}$ is distributed:

$$\sqrt{n}\left[\pi\left(\hat{\theta}\right) - \pi(\theta)\right] \sim N\left(0, \pi'\left(\hat{\theta}\right)\Sigma\ \pi'\left(\hat{\theta}\right)^t\right)$$

where $\theta^t = \left[\beta_1^t\ \beta_2^t\ \sigma_1\ \sigma_2\ \rho\right]$, Σ is the variance-covariance matrix of the parameters from the maximum likelihood function, while π' is a vector of partial derivatives with respect to θ evaluated at estimated coefficients $\hat{\theta}$.

The results in table 4 suggest that the proportion of consumers who are credit-constrained is 26 percent when the levels equation is estimated, and 31 percent when the equation system is estimated in log-levels. These figures is slightly larger than is usually estimated.

For instance, Hall and Mishkin (1982) estimate that around 20 percent of households are credit constrained, a figure that is close to the 19.4 percent estimated by Mariger (1987). Jappelli (1990) also estimates a figure of around 19 percent.[9] One explanation of why this study finds a higher figure than the self-reported responses contained in Jappelli (1990) is that this paper will also include those households who were allowed to hold some debt, but not as much as they wish to hold. His definition only included those households who were rejected outright.

This number was constructed under the assumption that the errors ε_{1i} and ε_{2i} in equation 1 are uncorrelated ($\rho = 0$). This assumption seems implausible. One interpretation of the error term is that it is due to parameters being omitted from the regression, either because they are unmodelled, or because they are not observed. For the parameters in the regression to be identified it must be true that $E(\varepsilon_{ji}|X_{ji}) = 0$ for $j = 1, 2$, the standard assumption in a regression. However, if $\sigma_{12} = 0$ then this argues that any such omitted variable can only enter either the supply equation, or the demand equation, but not both. This seems unlikely. If instead an omitted variable increased (or decreased) both supply and demand, then this would cause a positive correlation between the errors, in which case $\sigma_{12} > 0$, while if the omitted variable entered the supply and demand equations with opposite signs, then $\sigma_{12} < 0$.

B: *Differences across consumers.*

As highlighted in the introduction, one of the advantages of the approach taken in this paper is that differences across consumer types can be sensibly investigated. The first question that can be asked is are consumers with different observable characteristics differently credit constrained? However, the approach can go further than that: it can also investigate the *reason* for these differences across consumers. For instance, are these differences being driven by differences in the supply of loans, or by differences in the demand for loans?

Two objects are of interest; the level of credit constraints for households with a given characteristic, or the marginal effect of a given characteristic on the level of credit constraint holding everything else constant. The first is reported in table 4, and shows that single households headed by men are much less likely to be constrained than single households

[9] The different results may partly reflect the later period that is studied here.

headed by women, while households headed by a couple (at least when the equation is estimated in logs) are somewhere in between. Those households who did not complete school are much less likely to be constrained, as are black households, while the number of children makes little difference. In regressions C and D, the poorest households are much less likely to be constrained, only 14 percent of those whose income is around $1000. This figure rises until income reaches $6,400 and is then flat above this. Looking at age in regressions A and B shows that credit constraints fall by a third between the ages of 25 and 55. Regressions C and D divide the sample into those who completed at least two years of college and those who did not. For the poorly educated households, the age profile is approximately flat, while the tables shows there were large declines in credit constraints among the more highly educated households. Notice however that credit constraints nevertheless remain high among older households.

Table 5 reports the marginal effect of characteristics on credit constraints, while table 6 reports the median level of demand and of supply conditional on the X-variates. In calculating the effect of being male or female for instance, all other variables were fixed at their observed level in the data and then the variable of interest was set first to male, and then to female. And similarly for the other variables.

According to table 5 women are again much more likely to be constrained than men, but the effect is less dramatic than before. This is also true for the less well educated, for black households and when comparing across age groups. Interestingly, the very poorest households are the least likely to be constrained: the results in table 4 were because these households other characteristics made them very likely to be denied credit. Table 6 highlights that while the supply of credit is marginally lower, at most $120 lower in regression B, their demand is significantly higher, perhaps, as in regression C, by over $1,000. The table also shows that while supply is around $300 higher for couples, demand is higher than for men, and only marginally lower than for women. This contributes to the lower incidence of credit constraints among married couples. Regressions A and B show that the marginal incidence of credit constraints declines with age: table 6 shows that while supply falls gently with age, demand falls much more rapidly. Regression C shows that the slope of the supply curve is slightly steeper for college educated households, but that the slope of the demand curve, while fairly flat in both regression C and D for poorly educated households, for college educated households there is a large decline in demand between ages 25 and 55. This results in the incidence of credit constraints falling dramatically with age for these households. When differences in income are compared, table 5 shows that poor households are less likely to be constrained, but that beyond an income of $19,000 that the incidence of credit constraints are flat (and may even decline slightly). Table 6 shows that the demand for credit increases with income until it is around $19,000, and is especially low for the poorest households, supply is reasonably flat for lower income households, but increases rapidly for the richest households.

C: *The demand for debt among credit-constrained households*

The previous discussion investigated the extent of credit rationing among US households. However, a complete discussion will also consider how important rationing is for these consumers. In the absence of rationing how much more would households borrow? If the *extra* amount of debt that households wish to borrow is denoted Δ then the problem is to construct some estimate of[10]:

$$E\left(\Delta_i\right) = E\left(\Delta|\pi_i = 1\right) Pr\left(\pi_i = 1\right) + E\left(\Delta|\pi_i = 0\right) Pr\left(\pi_i = 0\right) \tag{8}$$

But by construction, the household can borrow as much as it likes if it is not credit constrained hence $\Delta_i = 0$ whenever $\pi_i = 0$, and this term drops out. Moreover, For the last term term $y_{2i} = y_i$ whenever $\pi_i = 1$, and y_i is observed, thus

$$E\left(\Delta_i\right) = Pr\left(\pi_i = 1\right)\left[E\left(y_{1i}|\pi_i = 1\right) - y_i\right] \tag{9}$$

The problem is to find some good proxy or estimate of the first part of equation 9. One approach, prevalent in the literature (see for instance Cox and Jappelli, 1993), is to replace $E\left(y_{1i}|\pi_1 = 1\right)$ by $E\left(y_{1i}|\pi_1 = 0\right)$. However, even if the estimate of $E\left(y_{1i}|\pi_1 = 0\right)$ is consistently estimated, using this in equation 9 will result in downward biased estimates of Δ_i.

[10]It should be understood that we are conditioning on all the observed characteristics of the agent, and on the observed level of borrowing, although the notation has suppressed this.

This is because (ignoring the zero observations).

$$\begin{aligned} E\left(y_{1i}|\pi_1 = 1\right) &= E\left(y_{1i}|y_{1i} > y_{2i}\right) \\ &> E\left(y_{1i}|y_{2i} > y_{1i}\right) \end{aligned}$$

A naive estimation strategy would thus under-estimate the true impact of credit-constraints on households. Instead construction of the difference entails allowing for the selectivity problem that this highlights. Given that f_1 and f_2 have both been estimated, then the expectation can be constructed:

$$E\left(\Delta_i\right) = \int_{y_i}^{\infty} \left(y_{1i} - y_i\right) Pr\left(y_{1i}|\pi_i = 1\right) dy_{1i}$$

where

$$Pr\left(y_{1i}|\pi_i = 1\right) = Pr\left(y_{1i}|y_{1i} > y_{2i} > 0\right) + Pr\left(y_{1i}|y_{1i} > 0; y_{2i} < 0\right)$$

and in the case of uncorrelated errors this becomes:

$$Pr\left(y_{1i}|\pi_i = 1\right) = \frac{\phi\left(\frac{y_{1i} - X_{1i}\beta_1}{\sigma_1}\right)}{\Phi\left(\frac{y_{1i} - X_{1i}\beta_1}{\sigma_1}\right)}$$

which can be recovered. Construction of this results in an estimate of $E\left(\Delta_i\right)$, see table 7, of \$1,655 dollars for the level regression, and around \$4,000 for the three log-level regressions. That is, in the absence of credit constraints, households would, on average, borrow over one and a half thousand more unsecured dollars than they currently do (or roughly \$4,000 for the log-level equations). This number is large given that the average amount of debt households hold is \$2,733.

The results also show that this problem is particularly acute for college educated households, and for married or female households. It is also much more serious for younger households, especially for college educated and younger households (by contrast age has little affect among non-college households). Poor and black households are much less seriously affected in terms of how much extra they wish to borrow (although this smaller figure may be a larger proportion of their permanent income and hence more serious in utility terms).

5 Conclusion

The paper described how it was possible to estimate a model in which some consumers were credit-constrained, even though it was not known which consumers were constrained. Estimation also addressed the selectivity problem: demand (or supply) conditional on being constrained does not equal the unconditional demand. Using the strategy proposed in the main body of the paper, the demand for debt and the supply of debt were separately identified and estimated. From this the paper recovered the incidence of credit-constraints among households, estimated to be between about 26 percent and about 31 percent in the population as a whole.

The paper found that sex, education, income, and race dramatically changed the incidence of credit-constraints, as well as age, at least for college educated households. In fact, the profile of a credit constrained household would be a single white female college graduate who has just started their first well-paid job. By contrast, a black male high school drop-out is far less likely to be constrained. Recall that poor and black households recorded a lower level and incidence of debt: this paper shows that this is overwhelmingly a demand, rather than a supply effect. One explanation is that the likely earnings profile of these households is flat, rather than increasing with age, and hence they do not wish to anticipate their future earnings. However, there is a steep decline in the demand for debt with age for college graduates. Two notes of caution should be made. First, poor households may be borrowing in markets that are not recorded in the CEX (for instance by borrowing through friends or using pawn dealers for instance). Secondly, the estimation strategy imposes that there are no fixed costs to asking for a loan (and hence no discouraged borrowers); if there were substantial fixed costs, then this would relatively reduce the estimated demand for small loans (typically given to poorer households). The justification for this important assumption is that modern credit scoring techniques, and the fact that almost all adults have a record with the major US credit bureaus, suggests that the costs to the bank of making a credit check on any potential customer is negligible (at most a few cents). Hence this paper claims this key assumption is reasonable.

Overall, the paper shows how three of the questions that were raised in the introduction can be answered and, in the appendix, it explains why it is not practically possible to answer the first of the four questions. The results are perhaps a little surprising, and show that credit constraints may be more pervasive than is commonly thought. Moreover, the paper also demonstrates that looking at the level and incidence of debt among various groups in society can be very misleading when investigating who is and who is not constrained. The incidence is not only higher among middle income, and college educated households, the degree to which the constraints reduce their borrowing is also much larger. However, although poorer households may not want to borrow much more, this extra amount may have more serious utility implications: the marginal utility of \$500 dollars to someone earning only \$1,000 may be much higher than the marginal utility of \$4,000 to someone earning \$50,000 dollars.

References

[1] Amel, D., (1993) "State laws affecting the geographic expansion of commercial banks", *Working Paper, Board of Governors of the Federal Reserve System.*

[2] Andrews, D. W. K., (1993) "Tests for parameter instability and structural change with unknown change point", *Econometrica*, 61, pp821-856.

[3] Barsky R. B., and Miron J. A., (1989), "The seasonal cycle and the business cycle," *Journal of Political Economy*, 97, pp503-534

[4] Carroll, C. D., and Summers, L. H., (1991), "Consumption growth parallels income growth: some new evidence", in Bernheim B. D., and Shoven J. D., eds. *National Saving and Economic Performance* University of Chicago Press, pp305-343.

[5] Cox, D., and Jappelli, T., (1993) "The effect of borrowing constraints on consumer liabilities", *Journal of Money, Credit and Banking*, 25, pp197-213.

[6] Duca, J. and Rosenthal, S., (1994), "Borrowing constraints and access to owner-occupied housing", *Regional Science and Urban Economics*, 24 (3), pp. 301-322.

[7] Eaton, J., and Gersovitz, M., (1981), "Debt with potential repudiation: theoretical and empirical analysis", *Review of Economic Studies*, 48, pp. 289-309.

[8] Garcia, R., Lusardi, A., and Ng, S., (1997), "Excess sensitivity and asymmetries in consumption: an empirical investigation", *Journal of Money, Credit, and Banking*, 29, pp165-176.

[9] Grant, C., (2000), "Bankruptcy, credit constraints, and insurance: some empirics", *CSEF working paper*, no. 40.

[10] Gropp, R., Scholz, J. K., and White, M. J., (1997), "Personal Bankruptcy and Credit Supply and Demand", *Quarterly Journal of Economics*, 112(1), pp217-251.

[11] Gross, D., and Souleles, N., (2002), "Do liquidity constraints and interest rates matter for consumer behaviour? Evidence from credit card data", *Quarterly Journal of Economics*, 114(1), pp149-185.

[12] Hajivassiliou, V., (1987), "The external debt repayments problems of LDC's: an econometric model based on panel data", *Journal of Econometrics*, 36, pp205-230.

[13] Hajivassiliou V. and Ioannides Y., (2002) "Dynamic Switching Regression Models of the Euler Equation with Imperfect Regime Clssofication", *LSE mimeo.*

[14] Hall, R. E., (1978), "Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence", *Journal of Political Economy*, 86, pp971-987.

[15] Hall, R. E., and Mishkin, F. S., (1982), "The sensitivity of consumption to transitory income: estimates from panel data on households", *Econometrica*, 50, pp461-481.

[16] Hartley, M. J., and Mallela, P., (1977) "The asymptotic properties of a maximum likelihood estimator for a model of markets in disequilibrium", *Econometrica*, 45, pp1205-1220.

[17] Hayashi, F., (1985), "The effect of liquidity constraints on consumption: a cross-sectional analysis", *Quarterly Journal of Economics*, 89, pp183-206.

[18] Hayashi, F., (1987), "Tests for liquidity constraints: a critical survey", in T. Bewley, ed., *Advances in Econometrics*, Fifth World Congress, vol. 2, Cambridge University Press.

[19] Hwang, H. -S., (1980), "A test of a disequilibrium model", *Journal of Econometrics*, 12, pp319-334.

[20] Jaffee, D., M., and Russel, T., (1976), "Imperfect information, uncertainty, and credit rationing", *Quarterly Journal of Economics*, 80, pp651-666.

[21] Jappelli, T., (1990), "Who is credit constrained in the US economy?" *Quarterly Journal of Economics*, 94, pp219-234.

[22] Jappelli, T., Pischke, S., and Souleles, N., (1998), "Testing for liquidity constraints in Euler equations with complementary data sources" *Review of Economics and Statistics*, 80(2), pp. 251-262.

[23] Johnson, Norman, L., and Kotz, Samuel, (1972), "Distributions in Statistics: Continuous Multivariate Distributions" *Wiley Series in Probability and Mathematical Statistics*, John Wiley & Sons Inc., New York.

[24] Kehoe, Timothy J., and Levine, David K., (1993), "Debt-Constrained Asset Markets," *Review of Economic Studies*, 60, pp865-888.

[25] Kocherlakota, Narayana, (1996), "Implications of efficient risk sharing without commitment", *Review of Economic Studies*, 63, pp595-609.

[26] Kooiman, P., van Dijk, H. K., and Thurik, A. R., (1985) "Likelihood diagnostics and Bayesian analysis of a micro-disequilibrium model for retail services," *Journal of Econometrics*, 29, pp323-354.

[27] Kroszner, R. S., and Strahan, P. E., (1999) "What drives deregulation? Economics and Politics of the relaxation of branching restrictions", *Quarterly Journal of Economics*, 114, pp1437-1467

[28] Mariger R. P., (1987), "A life-cycle consumption model with liquidity constraints: theory and empirical results," *Econometrica*, 55, pp533-557

[29] Miron J. A., (1986), "Financial Panics, the seasonality of the nominal interest rate, and the founding of the Fed" *American Economic Review*, 76, pp125-140

[30] Perraudin W. R.M., and Sørensen, B., (2000), "The demand for risky assets: sample selection and household portfolios", *Journal of Econometrics*, 97, pp. 117-144.

[31] Quandt, R., (1988), *The Econometrics of Disequilibrium*, Basil Blackwell, London.

[32] Stiglitz, J. E., and Weiss, A., (1981), "Credit rationing in markets with imperfect information", *American Economic Review*, 71, pp393-410.

[33] Zeldes, S., (1989), "Consumption and liquidity constraints: an empirical investigation", *Journal of Political Economy*, 97, pp305-346.

Appendix

The likelihood function

$$ln\mathcal{L} = \sum_i d_i \;\; ln\left(h_{1i}h_{2i} + h_{3i}h_{4i}\right) + \left(1 - d_i\right) ln\left(h_{5i}\right)$$

where d_i is an indicator function for observing a non-zero debt holding, that is:

$$d_i = I\left(y_i > 0\right)$$

and

$$h_{1i} = \left(2\pi\sigma_1^2\right)^{-\frac{1}{2}} \exp\left[-\frac{\left(y_i - \beta_1' X_{1i}\right)^2}{2\sigma_1^2}\right]$$

$$h_{2i} = \frac{1}{2\pi\left(1-\rho^2\right)^{\frac{1}{2}}} \int_0^\infty \int_{y_{1i}}^\infty exp\left[\frac{-1}{2\left(1-\rho^2\right)}\left(\xi_1^2 - 2\rho\xi_1\xi_2 + \xi_2^2\right)\right] d\xi_1 d\xi_2$$

$$h_{3i} = \left(2\pi\sigma_2^2\right)^{-\frac{1}{2}} \exp\left[-\frac{\left(y_i - \beta_2' X_{2i}\right)^2}{2\sigma_2^2}\right]$$

$$h_{4i} = \frac{1}{2\pi\left(1-\rho^2\right)^{\frac{1}{2}}} \int_0^\infty \int_{y_{2i}}^\infty exp\left[\frac{-1}{2\left(1-\rho^2\right)}\left(\xi_1^2 - 2\rho\xi_1\xi_2 + \xi_2^2\right)\right] d\xi_1 d\xi_2$$

$$h_{5i} = 1 - \frac{1}{2\pi\left(1-\rho^2\right)^{\frac{1}{2}}} \int_{k_{1i}}^\infty \int_{k_{2i}}^\infty exp\left[\frac{-1}{2\left(1-\rho^2\right)}\left(\xi_1^2 - 2\rho\xi_1\xi_2 + \xi_2^2\right)\right] d\xi_2 d\xi_1$$

and where $k_{1i} = \frac{\beta_1' X_{1i}}{\sigma_1}$, $k_{2i} = \frac{\beta_2' X_{2i}}{\sigma_2}$, and $\Phi\left(\cdot\right)$ is the usual *c.d.f.* of the normal distribution and the correlation of the errors is defined as $\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}$.

The intuition is straightforward: non-zero observations may either have been generated by the demand equation or by the supply equation. If generated by the demand equation, then h_{1i} represents the *p.d.f.* of the observation, given that it is generated by the demand equation, while h_{2i} represents the probability that demand is exceeded by supply. The converse is represented by h_{3i} and h_{4i}. Finally, the term h_{5i} is the standard bivariate normal c.d.f. and represents the probability that either supply or demand is less than or equal to zero.[11] If the correlation between the errors is identically equal to zero (in which case $\rho \equiv 0$) then h_{2i} and h_{4i} simplify in the obvious way, while h_{5i} becomes:

$$h_{5i} = 1 - \Phi\left(\frac{\beta_1 X_{1i}}{\sigma_1}\right)\Phi\left(\frac{\beta_2 X_{2i}}{\sigma_2}\right)$$

Results will be presented for this simpler framework.

[11] See Johnson and Kotz (1972) for a general discussion of the derivation of these equations.

Testing for credit constraints

It would be useful to test the system of equations in our model against an appropriate market clearing model. However, the following discussion will explain that this is not practically possible. In a market clearing model, the supply and demand equations jointly determine the level of borrowing, and there are no parameter restrictions implied by the model. Thus conventional tests are not appropriate, something that has been known at least since Hwang (1980).[12] To illustrate the argument, suppose, for the time being, the zero observations were ignored. As specified in equation 1 the model to be estimated takes the form[13]:

$$
\begin{aligned}
y_{1i} &= f_1\left(X_{1i}, \beta_1\right) + \varepsilon_{1i} & \quad & i \in \Omega_1 \\
y_{2i} &= f_2\left(X_{2i}, \beta_1\right) + \varepsilon_{2i} & \quad & i \in \Omega_2 \equiv \Omega \backslash \Omega_1
\end{aligned}
$$

where Ω_1 and Ω_2 partition observations between regimes and $\Omega_j \in \Theta$ where Θ is the space of all possible partitions. That is, Ω_2 represents those agents who are credit constrained, while Θ represents *all* the possible ways of selecting credit constrained households. By making appropriate assumptions about the parameters this system of equations can be estimated in a variety of ways, including FIML, SML, GMM, and MD. In the main part of the paper, the equations have been linearized, and then estimated by maximum likelihood. Estimation comprises both estimating $\hat{f}_1$ and $\hat{f}_2$, and estimating $\hat{\Omega}_i$. Having derived some estimate of f_1 and f_2 the challenge is to test the estimated model against some alternative model in which nobody is ever credit constrained (or agents are always credit constrained). In such a model the system of equations reduces to only one equation. That is, all observations will fall into only one of the regimes, f_1 say. The null and alternative hypotheses can thus be written:

$$
\begin{aligned}
H_0 &: \beta_i = \beta & \quad & \forall i \\
H_1 &: \beta_i = \begin{cases} \beta_1\left(\pi_i\right) & i \in \Omega_1 \\ \beta_2\left(\pi_i\right) & i \in \Omega_2 \end{cases}
\end{aligned}
$$

[12] He suggested a cusum or cusum of squares test on the residuals from the market clearing model but noted the poor power of the test in large samples.

[13] For the purposes of this discussion we will ignore the observations in which no borrowing occurs.

Let π_i denote an indicator function for observation i belonging to the first regime[14]: $\pi_i = I\left[y_{2i} - y_{1i} \geq 0\right]$. (And let π be the stacked vector of π_i's.) In which case the system of equations can be re-written in the following way.

$$y_i = f_{1i} + \pi_i \left(f_{2i} - f_{1i}\right) + \varepsilon_{1i} + \pi_i \left(\varepsilon_{2i} - \varepsilon_{1i}\right) \tag{10}$$

From this it is immediately apparent that there is a problem with testing the parameters of the model. Suppose that the aim was to test whether all observations were generated by the first equation (*i.e.* that credit constraints are never binding). Equivalently this can be interpreted as either π_i identically equals zero, or $f_{2i} - f_{1i}$ identically equals zero. The problem is that if $\Omega_2 = \phi$ (the empty set) then π_i and $f_{2i} - f_{1i}$ are not separately identifiable. If the null hypothesis had generated the data, then it becomes problematic to test against the alternative hypothesis.

A literature has recently developed which has explored ways of testing such models which are contaminated by nuisance parameters that only exist under the alternative hypothesis. If π where known, then testing the model would simply be a matter of constructing Wald, LR-, or LM- statistics and comparing the test statistic against a standard chi-squared distribution with degrees of freedom equal to the dimension of β. However, π is not known. Since it is derived in a way that is dependent on the data, the Wald (similarly LR-, and LM-) statistic based on this estimated π will no longer have a standard distribution: instead the null hypothesis will be *over-rejected*. This could lead to the mistaken conclusion that some agents suffer binding credit constraints. The existing literature (see for instance Andrews 1993) has developed tests for cases when the data is naturally ordered, but they can not practically be implemented here. To bootstrap the distribution of the Wald-statistic, as this literature suggests, would take an infeasible amount of time.

[14]This definition can be reconciled with the earlier definition of π if we note that now we are ignoring the zero observations.

Table 1: When US states deregulated and relaxed state banking regulations.

State	intrastate branching through M & A	full intrastate branching	State	intrastate branching through M & A	full intrastate branching
Alabama	1981	1990	Minnesota	1993	-
Alaska	1970	1970	Miss.	1986	1989
Arizona	1970	1970	Missouri	1990	1990
Arkansas	1994	-	Nebraska	1985	-
California	1970	1970	Nevada	1970	1970
Colorado	1991	-	New Hamp.	1987	1987
Connect.	1980	1988	New Jersey	1977	-
Delaware	1970	1970	New Mexico	1991	1991
D.C	1970	1970	New York	1976	1976
Florida	1988	1988	N. Carolina	1970	1970
Georgia	1983	-	Ohio	1979	1989
Hawaii	1986	1986	Oklahoma	1988	-
Idaho	1970	1970	Oregon	1985	1985
Illinois	1988	1993	Penn.	1982	1990
Indiana	1989	1991	S. Carolina	1970	1970
Iowa	-	-	S. Dakota	1970	1970
Kansas	1987	1990	Tenn.	1985	1990
Kentucky	1990	-	Texas	1988	1988
Louisiana	1988	1988	Utah	1981	1981
Maine	1975	1975	Vermount	1970	1970
Maryland	1970	1970	Virginia	1978	1987
Mass.	1984	1984	Washington	1985	1985
Michigan	1987	1988	W. Virginia	1987	1987
			Wisconsin	1990	1990

Taken from Amel (1993) and Kroszner and Strahan (1999), while 1970 means 1970 or before.

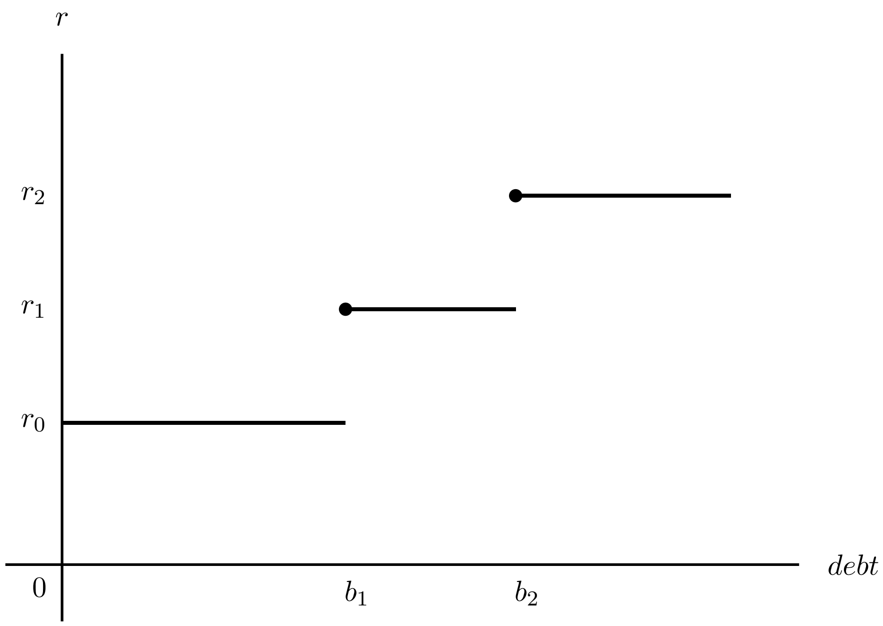

Figure 1: The marginal rate of interest as debt increases.

Table 2: Some summary statistics on debt-holding among US households.

	median debt ($)	ratio holding debt (%)	mean debt* ($)
All	736	68.0	3,984
1988	671	68.0	4,035
1989	865	71.9	3,748
1990	778	69.0	4,015
1991	777	68.4	4,192
1992	741	67.2	3,757
1993	554	64.1	4,176
Age 25-35	939	71.0	3,974
Age 35-45	732	67.7	4,301
Age 45-55	385	61.7	3,552
No Children	522	65.3	3,784
1 Child	1,019	71.5	4,249
2 Children	1,063	72.0	4,097
3-4 Children	939	69.1	4,362
some school	0	46.7	3,638
finished high school	657	67.3	3,388
some college	936	71.7	3,679
university degree	918	71.7	4,700
white	809	69.7	3,999
black	209	55.8	3,849
single male	164	55.8	3,750
single male	583	70.1	3,225
married	964	70.8	4,197
income $\leq$ 6,400	0	36.7	4,488
income 6,400-32,000	486	64.8	3,322
income $\geq$ 32,000	1,158	73.9	4,486

*Conditional on holding at least some debt. The table is constructed using the CEX.

Table 3: Estimated debt equations, in \$1,000's (standard errors in parenthesis).

parameter	A		B		C		D	
	demand	supply	demand	supply	demand	supply	demand	supply
constant	-7.496	1.824	-0.242	0.919	-0.084	0.895	-0.156	0.984
	(1.030)	(0.174)	(0.165)	(0.130)	(0.166)	(0.136)	(0.119)	(0.088)
female	2.456	-0.042	0.491	-0.045	0.492	-0.018	0.474	-
	(0.597)	(0.092)	(0.102)	(0.063)	(0.100)	(0.063)	(0.083)	
married	3.984	0.240	0.459	0.224	0.338	0.102	0.340	0.098
	(0.500)	(0.083)	(0.082)	(0.061)	(0.083)	(0.064)	(0.079)	(0.054)
age	-0.202	-0.055	-0.031	-0.031	-0.139	-0.008	-0.122	-0.051
	(0.449)	(0.073)	(0.073)	(0.053)	(0.117)	(0.091)	(0.103)	(0.053)
age^2	-0.323	-0.006	-0.014	-0.025	0.013	-0.034	0.013	-0.027
	(0.252)	(0.041)	(0.040)	(0.030)	(0.062)	(0.047)	(0.056)	(0.029)
age^3	-0.588	-0.019	-0.071	-0.007	0.053	-0.002	0.065	0.017
	(0.295)	(0.048)	(0.048)	(0.034)	(0.077)	(0.058)	(0.068)	(0.034)
High school	5.262	-0.028	0.596	0.068	0.493	0.055	0.510	-
	(0.570)	(0.095)	(0.091)	(0.079)	(0.093)	(0.081)	(0.078)	
Some college	6.274	0.086	0.739	0.136	0.613	0.044	0.625	-
	(0.586)	(0.072)	(0.094)	(0.079)	(0.113)	(0.095)	(0.098)	
College degree	7.155	0.017	0.746	0.167	0.573	0.054	0.587	-
	(0.542)	(0.093)	(0.087)	(0.077)	(0.109)	(0.095)	(0.093)	
age $\times$ College	-	-	-	-	0.203	-0.067	0.180	-
					(0.147)	(0.111)	(0.123)	
age^2 $\times$ College	-	-	-	-	-0.020	0.008	-0.015	-
					(0.078)	(0.059)	(0.065)	
age^3 $\times$ College	-	-	-	-	-0.218	-0.024	-0.243	-
					(0.098)	(0.073)	(0.082)	
non-white	-2.830	0.003	-0.370	-0.017	-0.302	-0.010	-0.306	-
	(0.445)	(0.080)	(0.071)	(0.062)	(0.072)	(0.063)	(0.061)	
one child	0.064	0.070	-0.019	0.071	-0.009	0.079	-0.000	0.063
	(0.517)	(0.080)	(0.082)	(0.061)	(0.078)	(0.062)	(0.078)	(0.062)
two children	-0.158	0.067	0.024	0.020	0.036	-0.015	0.047	-0.039
	(0.499)	(0.077)	(0.081)	(0.058)	(0.079)	(0.059)	(0.078)	(0.058)
3-4 children	-0.274	0.166	-0.031	0.059	0.003	0.061	0.011	0.041
	(0.610)	(0.094)	(0.097)	(0.075)	(0.095)	(0.076)	(0.094)	(0.075)
income	-	-	-	-	0.218	0.236	0.219	0.239
					(0.037)	(0.032)	(0.036)	(0.030)
income^2	-	-	-	-	-0.118	0.008	-0.114	0.088
					(0.022)	(0.021)	(0.022)	(0.020)
interest rate	-0.033	-0.008	0.004	-0.007	0.004	-0.012	0.028	-0.021
	(0.167)	(0.026)	(0.027)	(0.019)	(0.026)	(0.019)	(0.019)	(0.018)
law 1	-	-0.040	-	-0.001	-	-0.003	-	-0.001
		(0.091)		(0.058)		(0.059)		(0.058)
law 2	-	-0.062	-	-0.045	-	-0.037	-	-0.010
		(0.072)		(0.048)		(0.049)		(0.046)

All regressions included a constant, year, and the instruments (quarter in the demand equation; law 1 and law 2, allowing intra-state bank branching through mergers and acquisition and allowing full intra-state branching respectively, and people per bank in the state in the supply equation). Income is log-income, high-school refers to completed high school, 'age' refers to (age-40)/10, while the interest rate is the tax free municipal bond interest rate. Column A is estimated in levels, while the other columns are estimated in log-levels.

Table 4: The Incidence of credit constraints by type of household.

Subgroup	A		B		C		D	
	constr. (%)	*s.e.*	constr. (%)	*s.e.*	constr. (%)	*s.e.*	constr. (%)	*s.e.*
All	25.96	0.45	31.61	0.30	31.36	0.51	31.44	0.51
Men	19.29	0.85	25.13	0.97	25.65		25.74	0.82
Women	24.96	1.18	36.47	1.27	37.43	1.28	36.97	1.12
Married	28.00	0.51	32.05	0.47	31.6	1.00	31.79	0.48
Some School	12.28	0.87	18.35	1.06	19.80	1.26	18.92	0.80
Compl. school	24.84	0.77	31.07	0.79	31.45	0.82	31.02	0.67
Some College	27.17	0.79	33.63	0.75	34.05	0.77	34.49	0.66
College degree	29.75	0.69	33.73	0.61	32.56	0.57	33.03	0.52
White	26.99	0.48	32.55	0.51	32.36	0.52	32.44	0.51
Black	18.75	0.93	23.53	0.95	24.43	1.04	24.42	0.74
No children	24.24	0.57	30.91	0.65	31.04	0.67	31.01	0.66
1 child	28.70	0.92	31.46	0.74	30.86	0.72	31.06	0.71
2 children	28.14	0.86	33.40	0.76	33.10	0.76	33.42	0.76
3-4 children	26.73	1.13	30.22	0.93	30.47	0.97	30.44	0.95
Income - 1,100					14.50	1.47	14.10	1.40
Income - 6,400					24.87	1.05	24.79	1.01
Income - 19,000					32.91	0.69	32.93	0.68
Income - 54,000					32.46	0.42	32.63	0.42
Income - 145,000					32.02	0.41	32.18	0.41
Law 1 = no	27.35	0.67	33.67	0.70			31.43	0.59
Law 1 = yes	25.74	0.45	31.06	0.49			31.44	0.51
Law 2 = no	26.64	0.53	32.38	0.57			31.39	0.54
Law 2 = yes	25.65	0.46	30.99	0.50			31.46	0.52

In column A the regression is estimated in levels, while the other columns are estimated in log-levels. The regression results are displayed in table 3.

Table 4: cont. The Incidence of credit constraints by type of household.

Subgroup	A		B		C		D	
	constr. (%)	*s.e.*	constr. (%)	*s.e.*	constr. (%)	*s.e.*	constr. (%)	*s.e.*
Age 25	28.40	0.89	35.28	0.84				
Age 30	27.06	0.61	32.71	0.60				
Age 35	26.71	0.59	31.44	0.57				
Age 40	26.39	0.63	30.87	0.60				
Age 45	25.50	0.76	30.36	0.73				
Age 50	21.66	0.80	27.80	0.87				
Age 55	17.83	1.28	24.55	1.45				
No college education:								
Age 25					30.00	1.52	27.45	0.97
Age 30					31.00	0.99	29.47	0.73
Age 35					29.17	0.96	28.68	0.75
Age 40					26.51	0.96	26.49	0.77
Age 45					26.06	1.15	26.23	0.94
Age 50					25.96	1.28	26.89	1.16
Age 55					26.23	2.18	28.01	2.04
College education:								
Age 25					39.13	0.94	40.00	0.96
Age 30					33.81	0.66	34.76	0.65
Age 35					32.25	0.67	32.08	0.59
Age 40					31.74	0.62	32.35	0.62
Age 45					32.45	0.82	32.31	0.74
Age 50					30.01	1.11	29.21	0.89
Age 55					25.13	0.19	24.21	1.44
1988	27.00	0.99	31.41	0.92	31.41	0.94	32.70	0.66
1989	27.83	1.04	34.54	0.96	34.06	0.95	31.56	0.57
1990	27.03	0.98	32.70	0.97	33.17	1.00	32.11	0.57
1991	26.00	0.92	31.87	0.84	31.66	0.82	31.51	0.54
1992	25.23	0.85	30.77	0.82	30.53	0.82	31.32	0.55
1993	23.31	0.86	27.92	0.78	28.08	0.79	29.70	0.61

In column A the regression is estimated in levels, while the other columns are estimated in log-levels. The regression results are displayed in table 3.

Table 5: The Marginal Incidence of credit constrained households.

Subgroup	A		B		C		D	
	constr. (%)	*s.e.*	constr. (%)	*s.e.*	constr. (%)	*s.e.*	constr. (%)	*s.e.*
Men	19.55	0.98	27.98	1.12	28.38	1.05	28.49	0.91
Women	25.23	1.28	33.82	1.29	33.73	1.17	33.50	1.02
Married	27.73	0.55	31.64	0.50	31.44	0.58	31.58	0.58
Some High School	13.57	0.96	24.72	1.22	25.86	1.37	25.53	0.94
Compl. High school	25.00	0.78	31.27	0.80	31.15	0.90	31.31	0.71
Some College	26.94	0.79	32.29	0.74	32.43	0.83	32.48	0.67
College degree	29.10	0.68	31.18	0.60	31.97	0.74	32.10	0.61
White	26.71	0.47	31.84	0.51	31.72	0.53	31.81	0.52
Black	20.44	0.98	27.95	1.00	28.54	1.03	28.53	0.75
No children	26.18	0.64	31.49	0.62	31.37	0.64	31.36	0.62
1 child	26.11	0.98	30.90	0.88	30.83	0.87	31.00	0.87
2 children	25.63	0.93	31.87	0.93	31.84	0.93	32.09	0.93
3-4 children	25.12	1.20	30.84	1.10	31.06	1.11	31.23	1.11
Income - 1,100					22.14	1.83	22.04	1.77
Income - 6,400					29.57	1.19	29.64	1.17
Income - 19,000					31.99	0.73	32.10	0.72
Income - 54,000					31.19	0.46	31.26	0.44
Income - 145,000					29.18	1.04	29.24	1.00
Law 1 = no	25.86	0.49	31.42	0.58	31.34	0.60	31.43	0.59
Law 1 = yes	25.98	0.45	31.43	0.49	31.37	0.51	31.44	0.51
Law 2 = no	25.84	0.46	31.25	0.51	31.21	0.53	31.39	0.54
Law 2 = yes	26.02	0.45	31.51	0.51	31.44	053	31.46	0.52

These numbers are calculated by holding all other variables at their observed level, and changing the relevant variable to ‘male’, ‘female’ etc., and thus constructs the marginal effect of the relevant variable. In column A the regression is estimated in levels, while the other columns are estimated in log-levels. The regression results are displayed in table 3.

Subgroup	A		B		C		D	
	constr. (%)	*s.e.*	constr. (%)	*s.e.*	constr. (%)	*s.e.*	constr. (%)	*s.e.*
Age 25	28.94	1.26	33.50	1.10				
Age 30	26.79	0.61	31.93	0.60				
Age 35	26.09	0.63	31.28	0.62				
Age 40	25.96	0.63	31.12	0.62				
Age 45	25.49	0.79	30.96	0.76				
Age 50	23.72	0.86	30.27	0.87				
Age 55	19.70	1.78	28.40	1.88				
No college education:								
Age 25					31.91	2.27	30.49	1.49
Age 30					32.10	1.08	31.44	0.78
Age 35					31.65	0.86	31.41	0.71
Age 40					30.95	0.65	30.95	0.64
Age 45					30.40	0.95	30.64	0.86
Age 50					30.46	1.24	31.13	1.11
Age 55					31.61	2.81	33.17	2.53
College education:								
Age 25					34.22	1.18	35.10	1.12
Age 30					31.46	0.75	31.90	0.68
Age 35					30.61	0.70	30.76	0.66
Age 40					30.95	0.65	30.95	0.64
Age 45					31.35	0.88	31.24	0.82
Age 50					30.65	1.21	30.28	0.99
Age 55					27.26	2.98	26.27	2.20
1988	27.18	1.12	31.31	1.01	31.20	1.03	31.40	0.57
1989	27.72	1.07	32.70	0.96	32.44	0.96	31.22	0.53
1990	26.83	0.98	31.91	0.96	32.04	1.00	31.67	0.57
1991	25.82	0.92	31.58	0.84	31.52	0.83	31.45	0.55
1992	25.08	0.87	31.06	0.85	30.91	0.85	31.47	0.54
1993	23.76	1.00	30.13	0.93	30.27	0.93	31.42	0.56

Table 5: cont. The Marginal Incidence of credit constrained households.

These numbers are calculated by holding all other variables at their observed level, and changing the relevant variable to 'male', 'female' etc., and thus constructs the marginal effect of the relevant variable. In column A the regression is estimated in levels, while the other columns are estimated in log-levels. The regression results are displayed in table 3.

Table 6: Median demand and supply for credit.

Group	A		B		C		D	
	demand	supply	demand	supply	demand	supply	demand	supply
All	1,580	2,060	1,400	2,175	1,425	2,248	1,436	2,243
Men	906	1,904	638	1,729	791	2,023	802	2,021
Women	1,037	1,862	1,643	1,607	1,902	1,968	1,865	2,021
Married	2,138	2,145	1,561	2,417	1,491	2,349	1,512	2,333
Some High School	0	2,039	289	1,826	460	2,098	448	2,243
Compl. High school	1,000	2,010	1,269	2,025	1,349	2,276	1,366	2,243
Some College	1,677	2,125	1,616	2,241	1,646	2,237	1,654	2,243
College degree	2,363	2,056	1,633	2,342	1,543	2,270	1,554	2,243
White	1,745	2,059	1,489	2,182	1,498	2,252	1,510	2.243
Black	261	2,063	734	2,127	858	2,219	861	2,243
No children	1,618	2,019	1,402	2,130	1,409	2,192	1,409	2,220
1 child	1,661	2,089	1,357	2,360	1,387	2,456	1,408	2,430
2 children	1,514	2,087	1,462	2,066	1,497	2,144	1,525	2,095
3-4 children	1,439	2,186	1,329	2,320	1,418	2,393	1,435	2,356
Income - 1,100					90	1,377	91	1,399
Income - 6,400					698	1,344	701	1,337
Income - 19,000					1,348	1,733	1,358	1,717
Income - 54,000					1,607	2,764	1,620	2,772
Income - 145,000					1,313	5,125	1,319	5,252
Law 1 = no	1,580	2,095	1,400	2,180	1,425	2,259	1,436	2,246
Law 1 = yes	1,580	2,054	1,400	2,175	1,425	2,246	1,436	2,243
Law 2 = no	1,580	2,103	1,400	2,277	1,425	2,332	1,436	2,268
Law 2 = yes	1,580	2,040	1,400	2,132	1,425	2,211	1,436	2,232

In column A the regression is estimated in levels, while the other columns are estimated in log-levels. The regression results are displayed in table 3.

Table 6: cont. Median demand and supply for credit.

Group	A		B		C		D	
	demand	supply	demand	supply	demand	supply	demand	supply
Age 25	2,667	2,178	2,003	2,244				
Age 30	1,836	2,111	1,542	2,240				
Age 35	1,577	2,071	1,378	2,231				
Age 40	1,514	2,042	1,330	2,198				
Age 45	1,347	2,010	1,266	2,125				
Age 50	841	1,961	1,076	1,998				
Age 55	138	1,879	706	1,806				
No college education:								
Age 25					1,458	2,063	1,301	2,493
Age 30					1,558	2,166	1,481	2,380
Age 35					1,477	2,228	1,446	2,312
Age 40					1,320	2,241	1,314	2,243
Age 45					1,189	2,200	1,204	2,132
Age 50					1,161	2,100	1,218	1,944
Age 55					1,323	1,941	1,473	1,660
College education:								
Age 25					2,606	2,745	2,837	2,493
Age 30					1,544	2,498	1,600	2,380
Age 35					1,290	2,356	1,297	2,312
Age 40					1,320	2,241	1,314	2,243
Age 45					1,343	2,090	1,329	2,132
Age 50					1,087	1,852	1,056	1,944
Age 55					474	1,497	424	1,660
1988	1,889	1,998	1,380	2,198	1,387	2,257	1,436	2,265
1989	2,109	2,037	1,700	2,128	1,683	2,225	1,436	2,364
1990	1,824	2,046	1,473	2,092	1,507	2,092	1,436	2,126
1991	1,521	2,057	1,433	2,173	1,468	2,269	1,436	2,236
1992	1,338	2,096	1,314	2,174	1,331	2,276	1,436	2,226
1993	986	2,105	1,152	2,273	1,218	2,344	1,436	2,254

In column A the regression is estimated in levels, while the other columns are estimated in log-levels. The regression results are displayed in table 3.

Table 7: How much extra do households wish to borrow?

Group	A	B	C	D
All	1,655	4,036	3,882	3,916
Men	1,065	2,033	2,190	2,221
Women	1,550	4,782	5,173	5,067
Married	1,840	4,320	3,958	4,026
Some High School	643	1,167	1,398	1,350
Compl. High school	1,530	3,627	3,613	3,657
Some College	1,736	4,493	4,416	4,429
College degree	1,961	4,513	4,132	4,416
White	1,724	4,231	4,050	4,088
Black	1,147	2,250	2,370	2,372
No children	1,674	4,047	3,846	3,845
1 child	1,675	3,810	3,738	3,805
2 children	1,626	4,224	4,095	4,191
3-4 children	1,584	3,810	3,832	3,891
Income - 1,100			538	527
Income - 6,380			2,115	2,122
Income - 11,160			3,728	3,763
Income - 54,000			4,202	4,238
Income - 145,000			3,358	3,375

In column A the regression is estimated in levels, while the other columns are estimated in log-levels. The regression results are displayed in table 3.

Table 7: How much extra do households wish to borrow?

Group	A	B	C	D
Age 25	1,987	5,586		
Age 30	1,740	4,353		
Age 35	1,663	3,917		
Age 40	1,647	3,794		
Age 45	1,597	3,638		
Age 50	1,423	3,160		
Age 55	1,077	2,222		
No college education:				
Age 25			3,935	3,416
Age 30			4,179	3,909
Age 35			3,945	3,829
Age 40			3,526	3,491
Age 45			3,187	3,222
Age 50			3,134	3,298
Age 55			3,602	4,086
College education:				
Age 25			6,853	7,520
Age 30			4,072	4,226
Age 35			3,425	3,436
Age 40			3,526	3,491
Age 45			3,618	3,554
Age 50			2,992	2,871
Age 55			1,476	1,309
1988	1,771	3,952	3,755	3,912
1989	1,834	4,834	4,566	3,893
1990	1,740	4,228	4,119	3,941
1991	1,638	4,100	3,973	3,918
1992	1,569	3,779	3,601	3,920
1993	1,443	3,329	3,287	3,914

In column A the regression is estimated in levels, while the other columns are estimated in log-levels. The regression results are displayed in table 3.