The Free Installment Puzzle†

SungJin Cho
Seoul National University
John Rust
University Maryland

January, 2012

Abstract

We analyze a new dataset on borrowing decisions of a sample of customers of a credit card company. This credit card allows customers to pay for their purchases via *installment credit* over terms up to 12 months at an interest rate that depends on the customer's credit score and the duration of the installment loan. We use these data to estimate the effect of interest rates on consumers' demand for credit. We show that conventional econometric methods (including regression, instrumental variables, and matching estimators) predict that the demand for installment credit is an *increasing* function of the interest rate, an inference we dismiss as spurious due to the endogeneity of the interest rate and the effect of unobserved credit constraints that cause customers with worse credit scores to have higher demand for installment credit. To make more credible inferences about the effect of interest rates on the demand for credit we exploit a novel feature in our data: customers are more or less randomly offered *free installments*, i.e. the opportunity to pay back a given purchase over a fixed term ranging from 2 to 12 months at an interest rate of *zero*. We exploit these free installment offers as a *quasi-random experiment* the help identify the demand for credit by estimating a discrete choice model of the installment credit decision that accounts for censoring (choice based sampling) in observed free installments. Despite the significant censoring, we show that it is possible to identify consumers' choice probabilities and the probability they are offered free installments. The *free installment puzzle* results from our finding that less than 3% of the transactions in our sample were made as free installments, even though our model predicts that the average probability of being offered a free installment in our sample is approximately 20%. Our model predicts a high incidence of "pre-commitment behavior" even among the minority of individuals who do take the free installment offers. For example, the model predicts that 88% of individuals who were offered (and chose) a 10 month free installment offer pre-commited at time of purchase to pay the balance in *fewer* than 10 installments. This pre-commitment behavior is puzzling since there are no pre-payment penalties, and traditional economic models predict that consumers should choose the maximum loan duration when a loan is offered at a 0% interest rate. This puzzling consumer behavior raises questions about the company's behavior: why does it make so many free installment offers if the response to them is so poor? We also present evidence that the increasing interest rate schedule the company offers its customers may not be profit-maximizing.

---

†We are grateful to David McArthur for research assistance at an early stage of this project and for feedback from Ginger Zhe Jin and from participants at presentations of earlier versions of this paper at the University of Maryland in fall 2010, at the CEAR Workshop in Denver, January 2011, and at the State University of New York at Albany and Penn State University in September 2011. Please send comments to SungJin Cho at Department of Economics, Seoul National University, Seoul, Korea, email: `sungcho@snu.ac.kr` or John Rust at Department of Economics, University of Maryland, College Park, MD 20742 e-mail: `jrust@gemini.econ.umd.edu`.

1 Introduction

This paper presents new findings on the demand for credit based on a unique data set that allows us to observe borrowing decisions made by a sample of customers of a major credit card company. Unlike traditional *revolving credit* provided by most U.S.-based credit cards, the main type of credit contract offered by the company we study is *installment credit.* This is a common contract used by credit card companies in Latin American countries. Installment credit contracts require customers to make *ex ante* choices of the number of installments over which they will pay back the amount of each purchase, and they do this on a *transaction by transaction basis.* Customers are aware that they have this opportunity because it is described to them on each of their monthly statements, along with the interest rate schedule that determines the interest rate they would pay for installment loans payable over to 2 to 12 billing statements (months).

In contrast, under revolving credit customers do not make borrowing decisions on a transaction by transaction basis. Instead, their borrowing decisions are made at they time they pay *each bill.* Revolving credit amounts to an option pay only part of their balance due, and to use a sequence of one period loans of endogenously chosen sizes (subject to an overall credit limit) that allows customers to pay off their balances according to their own chosen time path. The company we study did not offer revolving credit to most of its customers until 2005, and then only to a minority of its customers with the best credit scores. Thus, without access to revolving credit, a customer's entire credit card balance is due and payable at each statement date unless the customer chose to pay for some of their purchases on installment.

A credit card company provided us with data on all purchases, billing statements, and payments made by a sample of 938 of its customers from late 2003 to spring 2007. We observe over *180,000 individual purchase transactions* for these customers over this period, and in the vast majority of these transactions constituted *micro-borrowing decisions* about whether to pay for the purchased amount in full at the next billing statement (which we denote as the choice $d = 1$) or to make the purchase under installment credit over 2 to 12 subsequent billing statements (which we denote as a choice d in the set $\{2,\dots,12\}$).

To our knowledge there is no previous study that analyzes these sorts of micro-borrowing decisions, especially at the level of detail and with the huge number of observations that we access to in this data set. In addition to having considerable data on the amount and type of the transaction, we also observe the company's proprietary credit scores for these customers, and we resolved problems of unobserved pre-sample balances (initial conditions) and are able to recreate the trajectories of their credit card and installment balances. We are also able to uncover (econometrically) the formula the company uses for setting installment credit interest rates, and we show that these interest rates not only depend on the credit score of the customer, but also on the duration of the installment loan. We show that the credit card company uses a particular non-linear increasing interest rate schedule that is *common* to all its customers. Thus, while the intercept of the interest rate schedule does shift to reflect consumer credit score and other credit history information, the schedule of interest rates for installment loans above a "base rate" for 2 month loans is common to all customers. So, for example, the interest rate the company charges for a 12 month installment loan is 7 percentage points higher than the interest rate it charges to a customer for a 2 month installment loan and this differential is the same for all customers.

The main goal of this paper is to use these data to infer the *credit demand function* and determine its elasticity with respect to the interest rate charged. Unfortunately, we show that conventional reduced-form econometric approaches, including regression, instrumental variables, and matching estimators, all imply that the demand for credit is an *upward sloping* function of the interest rate charged to consumers. Of course, we believe this is a spurious finding, a likely result of unobserved factors that make consumers who have high need for credit to be charged higher interest rates than consumers who have better credit scores or other lower cost borrowing opportunities, or who are otherwise not "liquidity constrained." Though we have reasonable instrumental variables (such as the Certificate of Deposit or "CD rate") that lead to credible, exogenous variation in the company's cost of credit (and therefore we presume exogenous variation in the interest rates it offers to its customers), in practice the "markup" the company charges to its customers over this CD rate is huge and highly variable and much more responsive to other factors such as credit card competition than it is to the relatively minor variations in the cost of credit to banks. As result we find that the CD rate and other similar instrumental variables are actually very *weak instruments* that are nearly uncorrelated with actual interest rates the company charges its customers. To the extent there is any correlation at all, we find customer interest rates are slightly *negatively* correlated with the CD rate and other similar instruments!

To make more accurate inferences about the demand for credit, we estimate a discrete choice model of a consumer's choice of installment loan duration (i.e. the choice of the number of installments d over which the amount purchased is paid back). The model has a flexible specification, so depending on the value of its parameters, it can approximate a wide variety of rational as well as "behavioral" theories of decision making. The model also accounts for the increasing, time-varying and customer-specific interest rate schedules that are difficult to handle using conventional regression methods. Most importantly, it also enables us to exploit the quasi-random variability in the interest rates charged to consumers as a result of *interest-free installment opportunities* that arise from promotions offered by the credit card company, sometimes in conjunction with merchants. We treat these free installment offers as *quasi random experiments* because executives at the credit card company assured us that the chance of being offered a free installment does not directly depend on customer characteristics, or the amount of their purchase. There may be *indirect selection effects* if customers are more likely to shop at a store and make purchases there if they know that they can take advantage of a free installment offer at that store over a particular time interval. However we believe the company executives that the mechanism by which free installments are offered to its customers does not directly depend on their charcteristics, or the amount of their purchase.

However we also confront econometric problems due to significant *censoring* (choice-based sampling) in free installment offers. That is, we only observe a subset of free installment offers that customers actually chose: we do not observe offers that were made and not chosen. Further, the company provided us with no data to independently estimate the probability distribution of how free installment offers were provided to customers over time and across different merchants. Despite the econometric challenges (we show that accounting for censoring results in a likelihood function that is akin to a mixture of choice probabilities, making our model potentially difficult to identify) we show that the conditional probability of free installment offers can be separately identified from customers' choice probabilities, and that we can even identify the probability distribution of the maximum duration of different free installment offers. We exploit the *a priori* information from the company executives that free installments do not depend on customer characteristics (or the purchase amount) as a powerful *exclusion restriction* to help identify our model. We show that our estimated model provides remarkably good predictions of the borrowing decisions of our sample of consumers, and can successfully control for the endogeneity of interest rates, resulting in a downward sloping demand for credit.

We find that the demand for credit is highly inelastic and the take up rate for free installment offers is surprisingly low: we estimate that on average, the probability that customers who are offered free installment opportunities will actually take them is only 15%. Instead, in the vast majority of cases, customers choose to pay the purchased amount in full at the next statement date. Of course, our model predicts that the probability of purchasing under installment is higher the larger the amount paid for a given transaction, and individuals who we suspect are "liquidity constrained" are uniformly more likely to take advantage of free installment offers than individuals who do not appear to be liquidity constrained.

Our estimated model leads to an even more puzzling prediction: a large fraction of the customers who are offered and actually choose free installment offers engage in *pre-commitment behavior* in the sense of making an *ex ante* decision to pay off their purchase in *fewer installments* than the maximal number of installments allowed under the free installment offer. For example, our model predicts that 88% of individuals who were offered and who chose a 10 month free installment offer pre-commited at the time of purchase to pay off their balance due in fewer than 10 installments. This pre-commitment behavior is puzzling since there is no pre-payment penalty in installment loans, so traditional economic theories predict that, barring special explanations to the contrary, rational consumers should never pre-commit to a free installment offer for a term that is less than the maximum offered. Doing so is to arbitrarily limit their future options without receiving any obvious *ex ante* compensation for doing so. We find that only a small minority of customers who are offered free installment loans would choose the maximum installment term offered to them (fewer than 1% of those offered 12 month loans, slightly over 2% of those offered 10 month loans, and approximately 10% of those offered 3 month free installment loans). The apparent aversion these customers have to taking advantage of zero interest loan opportunities constitutes what we call *the free installment puzzle.*

This aversion is very hard to explain using the standard economic model of behavior by rational individuals who maximize the expected discounted value of a time-additive utility with geometric discounting of future utilities. Early work by Strotz [1955] and subsequent contributions by Laibson [1997] and Gul and Pesendorfer [2001] and others on hyperbolic discounting, temptation, and self-control have shown that time-inconsistent behavior can arise in variety of extensions of the standard model of time-separable geometrically discounted utility maximization. Versions of these theories for "sophisticated" agents (i.e. agents who are self-aware of their time-inconsistent behavior) can explain a desire by some of these individuals to pre-commit to actions that restrain the options available to their "future selves". As Gul and Pesendorfer [2001] note, there are situations where pre-commitment can make these individuals "unambiguously better off when ex ante undesirable temptations are no longer available" (p. 1406).

Casari [2009] notes that "Although the implications of naïveté or sophistication are profound, the behavioral evidence is still quite limited" (p. 119). However there is some evidence, including laboratory evidence that Casari provides in his paper, that "the demand for commitment was substantial" even though

“Commitment always carries an implicit cost due to the uncertainty of the future.” (p. 138).[1]

Our findings are also puzzling in view of the conventional wisdom that many credit card customers are liquidity constrained and willing to borrow at usuriously high rates of interest. Indeed, at the same time as we infer large fractions of the customers in our sample forgoing free installment opportunities, other customers are paying very high rates of interest, averaging about 15%, to borrow varying amounts over varying lengths of time under traditional positive interest installment purchases. Our results are also puzzling in view of the aggressive use of free installments by credit card companies as a marketing tool in an attempt to gain a larger share of the credit card market. Why do these companies use free installments so frequently if they are aware that the take up rates of free installment offers are so low?

Finally, the highly inelastic demand response that we find to variations in interest rates is a puzzle, since we would expect that especially individuals who are liquidity constrained would have a strong motivation to use free installment credit opportunities at nearly every opportunity that they are offered to them. Although we have no precise way of identifying customers in our sample who are liquidity constrained, there is substantial heterogeneity in the free installment take up rates in the customers in our sample. We tentatively identify the individuals with the highest take up rates as those who are potentially liquidity constrained, though some of them could also be the rational time-separable, geometric discounted expected utility maximizers — i.e. *homos economicus* — who are predicted to ruthlessly exploit every free installment opportunity that is presented to them.

Section 2 describes the credit card data and documents the importance of merchant fees as a significant component of the profit that this company earns: we believe this is the main motivation for the company’s frequent use of free installments. Section 3 introduces the econometric methods we employed to infer the demand for credit starting with the more traditional regression-based and reduced-form treatment effect approaches. We show that the empirical findings from these reduced form methods result in implausible estimates of the demand for credit. In particular, all of the methods lead to the conclusion that the demand for credit is an *increasing* function of the interest rate.

Section 4 introduces our discrete choice econometric model of installment choice and derives the like-

[1]Ashraf et al. [2006] find a correlation between hyperbolic types of discounting and choice of a commitment savings product from a randomized controlled field experiment in the Phillipines. They note that “identifying hyperbolic preferences and observing a preference for commitment is difficult” but by “Using hypothetical survey questions, we identify individuals who exhibit impatience over near-term trade-offs bt patience over future trade-offs. Although we find this reversal uncorrelated with most demographic and economic characteristics, we do find thaat this reveral predicts take-up of a commitment savings product, particularly for women.” (p. 668).

lihood function for the discrete choice model accounting for the censored, choice-based nature of our observations of free installment offers (i.e. that we only observe free installment offers when customers actually choose them, not when customers do not choose them). We establish the identification of the structural parameters, and present the estimation results, including an evaluation of the goodness of fit of the model and the predicted installment credit demand function, as well as several counterfactual predictions of customer response to alternative installment credit policies. In particular, using the estimated demand system we search for alternative *consumer-specific* interest rate schedules that result in higher profits to the credit card company subject to the constraint that the expected utility of this alternative schedule to the customer is no lower than their utility under the company's current or *status quo* interest schedule. Our calculated optimal interest rate schedules differ significantly depending on customer characteristics and generally are very different from the particular schedule that the company has chosen. We view this as a further puzzle raised by our analysis.

Section 5 presents our conclusions and speculative comments about the underlying reasons for the free installment puzzle, as well as suggestions for future research provided additional data and particularly new experimental data could be gathered.

2 Credit Card Data

Our data consist of six data files: sales, billing, revolving and collection, credit rating, and a final file defining the merchant classification codes that appear in the sales data. For sales data, we should note that there are three types of sales 1) sales payable in full at the next statement date, 2) sales payable in installments over two or more statement dates, and 3) cash advances. Cash advances can either be paid in full at the next statement date, or paid by installment over multiple future statements. Generally purchases and cash advances that are paid by installment are done at relatively high interest rates, except when customers are offered free installment options.

We observe installment purchases of varying lengths, from 2 to 12 months. The most commonly chosen term is 3 months: 61.5% of all of the installment purchases we observe have a 3 month term. The maximum installment term we observe is 12 months, and is chosen in 1.7% of the cases. Other frequently chosen terms are 2 months (20.0% of cases), 5 months (5.0%), 6 months (4.9%), and 10 months (3.7%). There are no installment purchases with a term of 1 month, since this is equivalent to a regular charge, i.e.

a payment due at the next billing statement. Thus, we define the "installment choice set" for a consumer as being $D = \{1, 2, \ldots, 12\}$ where a choice of $d = 1$ is equivalent to a regular charge that will be due at the next billing statement, a choice of $d = 2$ corresponds to equal installments payable in the next two billing statements, and so forth, so that $d = 12$ denotes an installment contract that is payable over the next 12 billing statements (which typically arrive monthly).

Customers typically pay off their installment purchases in equal installment amounts. For example, if a consumer purchases an amount P under an installment contract with a total of d installments payments, then the consumer will pay back the "principal" P in d equal installments of P/d over the next d billing periods. If the consumer is charged interest for this installment purchase, the credit card company levies additional interest charges that are due and payable along with the installment payment at each of the successive d statement dates. However in some cases there are unequal payments, sometimes as a result of late payments, or accelerated or pre-payment of installments. The installment agreement does not formally allow for a pre-payment option, so that if a consumer does pre-pay an installment contract, the credit card company still charges the interest at the successive d statement dates, as if the customer had not pre-paid.

We calculate the realized internal rate of returns for 8987 installment transactions in our credit card data set. The internal rate of return is the interest rate r that sets the net present value of the stream of cash flows involved in the installment transaction to 0, where the initial purchase is regarded as a cash outflow (from the credit card company) at time $t = 0$, and the successive payments (including interest) are treated as cash inflows at the successive statement dates $t_1, t_2, \ldots, t_d$. There were only 141 cases out of the 8987 installment transactions where the customer did not follow the original installment contract by paying in the d installments that the customer originally agreed to pay. There were pre-payments in 127 cases, i.e. where the customer paid off the installment balance more quickly than necessary under the original installment agreement. Given that there is no direct benefit to the customer from pre-paying the installment (since the credit card company will continue to collect interest from the customer as if the installment loan had not been pre-paid), it seems hard to rationalize these cases under a standard model of a rational, well-informed consumer. In 31 of these cases, the customer was given a 0% installment loan, and yet still pre-paid. One possible explanation is that these customers were not aware that they had what was in effect an interest-free loan, and not aware that there was no benefit to pre-paying. These customers might have believed (incorrectly) that by paying off their installment balance more quickly they were saving interest charges, or perhaps some other explanation such as "mental accounting" (e.g. the desire to be free of the mental burden of having a large outstanding installment balance to pay), that might explain this behavior.

There were only 17 cases where the number of installment payments were greater than the number of installments originally agreed to in the original installment transactions. These do not appear to be "defaults" since the total amount collected in each of these cases equals the initial amount purchase. The delay in payment was typically only one billing cycle more than the originally agreed number of installments. For this reason, we believe that these cases might reflect the effect of holidays (such as where a payment is allowed to be skipped since a statement falls on a special holiday) or some other reason (e.g. an agreed *ex post* modification in the installment agreement). Since there are so few of these cases, we basically ignore them in the analysis below.

In the data we observe most installment purchases have a positive internal rate of return, but in nearly half of all installment purchases we observed (47.7%) the internal rate of return was 0, so the customers were in effect given an interest-free loan by the credit card company. These "zero interest installments" are usually a result of special promotions that are provided either at the level of individual merchants (via agreement with the credit card company to help promote sales at particular merchants via the "free credit" aspect of an installment purchase with a 0% interest rate), or via "general offers" that the credit card company offers to selected customers during specific periods of time either to encourage more spending, increased customer loyalty, or as a promotion to attract new customers. Our data does not contain enough information for us to determine exactly which customers are offered 0% installment options, so we model them as occurring probabilistically, depending on the merchant code where the customer makes a purchase, and dummies for the date of purchase (since some of these promotions tend to be offered at specific times in the year). The vast majority of interest-free installment loans have a term of 6 months or less. If a customer wishes to have a longer term than the one being offered, the customer generally must pay a positive interest rate for longer term installments, according to the schedule described below. In our analysis below, we will assume that when a customer is offered a interest-free installment purchase option, the maximum term is exogenously specified according to a probability distribution that we will estimate from our data.

In order to make customer-specific profit and rate of return calculations and analyze time patterns of credit card spending and installment usage, we had to assemble the data that were contained on customers in the sales, billing, and collections tables into a *longitudinal format* that would enable us to track the evolution of both credit card and installment balances on a *day by day basis*. We emphasize that the credit card company did not provide us with these latter data, rather we had to *construct the longitudinal data from the information we were provided.* While at first it may seem to be a relatively trivial exercise in stock/flow accounting to reconstruct these *balance histories* from the sales, billing and collection data, we faced a significant *initial conditions problem.* That is, we were not given the outstanding installment and credit card balances at any initial date. Instead the collections table would tell us the *statement amount* and information on dates of collection and amounts received, but without knowing an initial balance, it was not always easy to determine if a customer had paid the initial statement or any previous statements in full, or had unpaid balances that needed to be carried over from previous statement dates. We could obtain some indirect evidence of the presence of such overdue balances from late fees charged, but without going into more detail, it proved to be a rather challenging accounting exercise to infer the initial balances of the customers in our sample accounting for the variable left and right censoring in the data.

In particular, not all sales records in the sales table could be matched with billing records in the billing table and vice versa. In some cases, we observed purchases that were at a date before any date in the billing table, and we also observed billing records for which we could not find a corresponding record in the sales table. Fortunately the billing table had redundant information on whether the transaction was on installment or not, so in most cases we could reconstruct an entire installment transaction even if we only observed a truncated series of installment payments in the billing record and no record of the initial sale in the sales table.

Similarly there were also problems of right censoring in our data, since in many cases we observe sales in April 2007 for which we had no corresponding billing records, or no collection records at the end of a balance history that would enable us to determine whether an outstanding balance would be fully paid at the next (yet to be observed) statement date that was missing in the collection table. In such cases after making the best inference on the value of the customer's initial balance at the start of the interval we observed the customer, we followed the customer for as long as possible so we could also match every sale with its corresponding record in the billing table and track payments received on balances due in the collections table. In some cases this required us to "back up" by one or more months on the full history of the customer and discard transactions in the last month when we could not find matching records in the billing table and a record of payment in the collections table.

However, overall, our care in preparing the data paid off and we did not lose too many observations by doing this and the result is a considerably more accurate record for making profit/loss calculations on a customer by customer basis. If we did not do this, customers would be artificially classified as being in deficit if a balance due happened not to have been recorded for them in the collections table due to right censoring. Thus, we would end a record on a customer on a date where a balance due was received and for which all previous charges up to that date had been accounted for. Any subsequent charges that were made by the customer that would be billed and paid for in the future but which we could not yet observe in the billing or collections tables were discarded in our analyses of customer level profitability and returns.

Figure 1 plots our constructed longitudinal balance histories for one of the customers in our data set. We chose this example because the customer made only a single installment transaction and this makes it very easy to understand how the constructed balance histories behave. The top left panel of figure 1 is the overall creditcard balance for this customer. We start observing this customer making a charge of $118.30 on December 12, 2003. However we did not know what the outstanding balance was for this customer at this date since the first statement date for the customers was on January 20, 2004. We were able to determine in this case that this customer had no outstanding unpaid balances and we were able to allocate all charges the customer made in the sales table to matching entries in the billing table and thus track this customer with an accurate determination of the customer's initial balance at the first installment date. Thus, the top right panel of figure 1 displays our inferred balance for this customer, $427.24, on the first statement date we observe for this customer, January 20, 2004.

The dashed vertical lines in the figures represent the statement dates. Because this company has links to its customers' bank accounts and auto-debits the amount due on each statement date, its customers almost always pay the full balance due *exactly* on each statement date, unlike for many American credit card companies where customers may mail in a check or pay online and the date paid may often be plus or minus the statement date by several days. Thus, this feature leads to the inverted sawtooth appearance of balances in the top right hand panel of figure 1: balances tend to grow monotonically (though stochastically) between successive statement dates representing the spending the customer is doing on their credit card, then it drops discontinuously on each statement date representing the payment of the balance due.

Note that the discontinuous drops in the credit card balance at each statement date do not bring balances exactly to zero. The reason is that the credit card company assigns to each purchase a particular statement date at which that purchase will be due (unless it is an installment, which leads to a different treatment we will discuss shortly) and therefore any purchases a customer makes that are sufficiently close to an upcoming statement date will be assigned as due and payable by the company to the *following* statement

Figure 1: Balance and credit history of customer 125

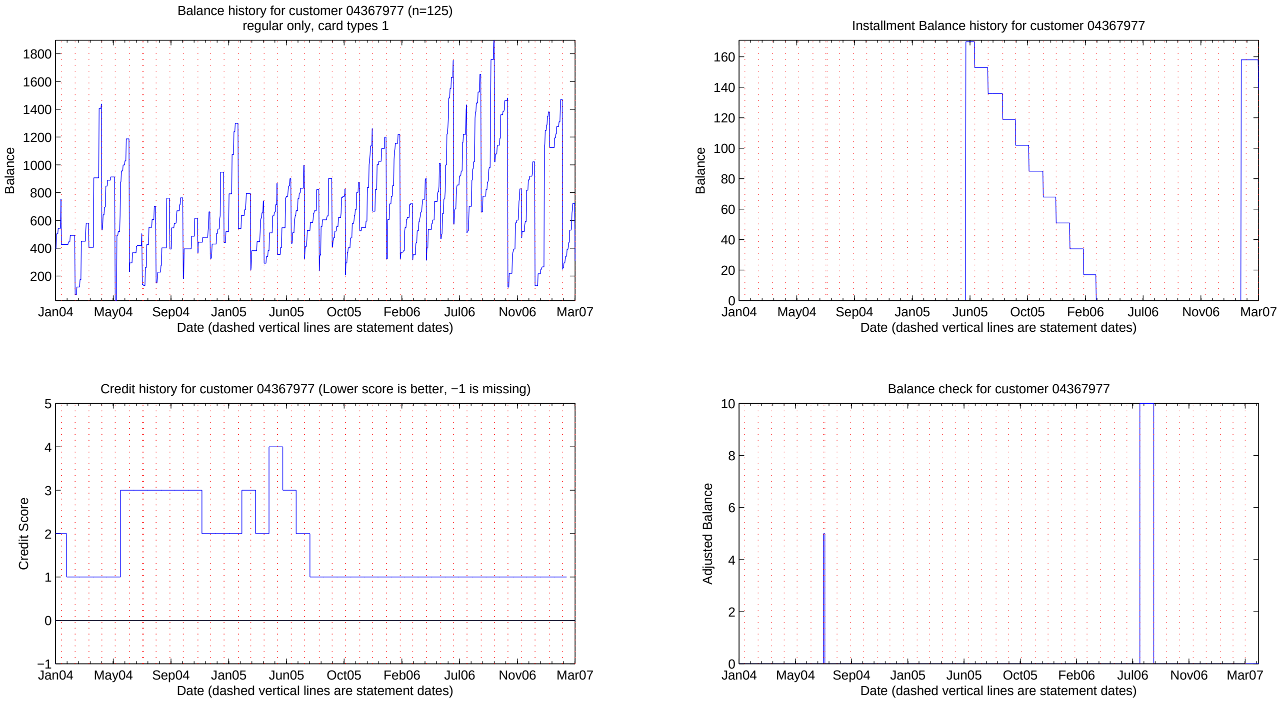
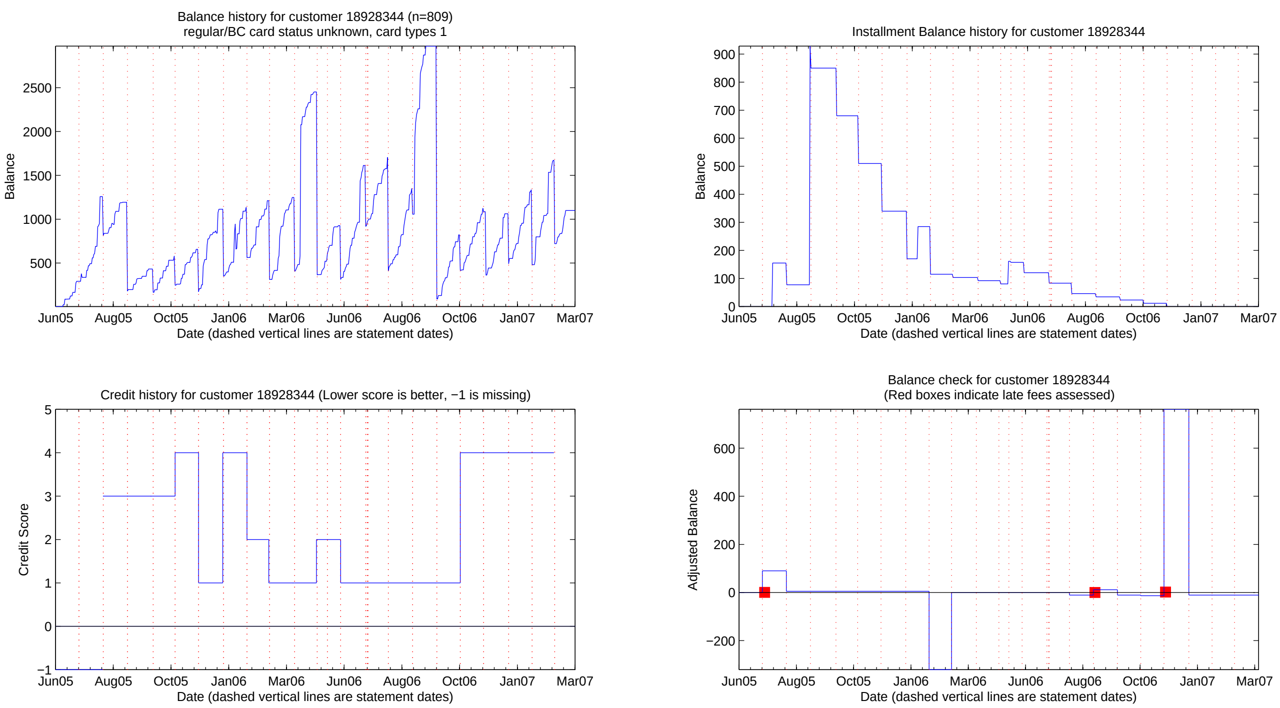


Figure 2: Balance and credit history of customer 809



date. Thus, the level of credit card balances just after a statement date reflects the sum of all purchases made prior to that statement date that the company assigned to be due and payable at the next statement date. This implies that a person's credit card balance will almost never be exactly zero, even on a statement date — at least for customers who are sufficiently active users of their credit card.

Note the "balance check" in the lower right panel of figure 1. The balance check should be identically zero if we had correctly inferred the customer's initial balance and perfectly tracked all charges and fees. However there were some small charges and payments that we could not reconcile or ascribe to any late charge, annual fee or so forth. These appear as the spikes in the lower right panel of figure 1. In some cases the balance check will be non zero due to a pre-payment or some slightly mis-timed or out of sync payment but shortly after the balance check returns to zero showing that we have basically correctly calculated the full balance history for this customer.

Now consider the top right panel of figure 1, which shows the *installment balance history* for the customer. We keep two separate accounts for the customer, 1) the credit card balance and 2) the installment balance. In this case, we see that the customer did not charge anything on installment until May 31, 2005 when the customer made an installment purchase in the amount of $169.90. This is reflected by the discontinuous upward jump in the installment balance in the top right panel of figure 1. We can see from the graph that this balance was paid off in 10 equal installments of $16.99. This installment also happened to be an interest-free installment and so at each of the 10 succeeding statement dates after the item was purchased on May 31, 2005 the installment balance decreased by $16.99 until the balance was entirely paid off at the statement date of March 20, 2006. Note that on each such statement date, the amount currently due on the customer's installment balance *transfers* and is added to the customer's credit card balance.

The final, lower left panel of figure 1 plots the credit score that the company maintained on this customer. Credit scores are integers on a scale from 1 to 10 with 1 being the best possible credit score and 10 being the worst. This customer generally had excellent credit scores, though for reasons that are not entirely clear from figure 1, the customer had periods of time (particularly May to September 2004 and May to July 2005) where the customer's credit score deteriorated for some reason. We see that the customer's worst credit scores appear to have coincided with the customer's installment purchase in May 2005.

We present another balance history for a more interesting customer, customer 809, in figure 2. This customer generally maintained larger credit card balances and also larger installment balances than customer 125, and we see that this customer also tends to have uniformly worse credit scores than customer

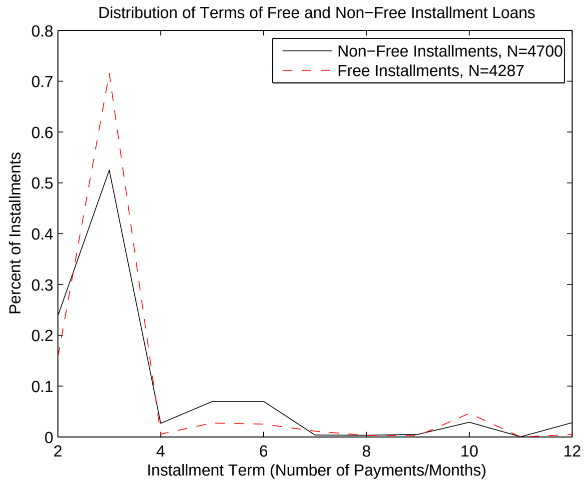
125 had. The red boxes in the lower right panel of figure 2 also indicate another behavior that is a big "no-no" for the credit card company: the customer was late in making payments and assessed late payments on three occasions. Because balances due are automatically debited from the customer's bank account, this means that on these three occasions the customer's bank account was *overdrawn* and the credit card company was unable to collect the full statement amount due. While the customer may have also been charged penalties by his/her bank, the late payment penalties charged by this credit card on these three occasions were trivially small by American standards: $0.18 in each case. The main penalty seems to be a degradation of the credit score, though the late fee of $0.45 that the customer was assessed on September 4, 2006 did not seem to have any effect on the credit score around that time.

Now that we have shown how we were able to construct the spending and payment patterns and thus the balances histories of our sample of customers dynamically, we are now in a position to calculate returns and profitability on a *customer by customer basis.* In terms of profits, we can think of the primary cost of a customer is the company's *cost of credit,* i.e. the credit card company's borrowing cost or opportunity cost of capital. In the case of customers who default, the company also loses the unpaid balance of their loan to the customer. The revenues include annual fees, late fees, interest and service charges, and merchant fees. We note that our measure is one of *gross profits*, i.e. we do not know the cost of things such as 1) rewards programs, 2) advertising costs, and 3) other fixed operating costs such as billing and collection costs and wages and salaries and payments to other credit card companies for out of network transactions.

Figure 3 plots the distributions of installment terms for 4700 installment transactions made by customers that chose installment with positive interest rates, and also the distribution of installment terms offered to 4287 customers who chose free installment offers. The distributions are roughly similar except that the mean installment term chosen by customers under positive interest installments, 3.66 payments/months, is longer than the 3.42 payments/months offered to customers who chose free installment options. We see that when customers choose installments with a positive interest rate, they are generally more likely to choose longer payment terms, though the difference in the two distributions is not particularly striking.

Note that due to censoring we are not always able to observe the full duration of installment transactions. For example we observe some installment transactions in our billing data for which the date of the initial installment purchase is not in our sales table. This is why, although we can identify 11175 installment transactions in our billing data, when we eliminate censored observations we obtain a smaller set of

Figure 3: Durations of Free and Non-Free Installment Loans



8987 *uncensored* observations of installments where we can match the transaction in the billing table to the original sale in the sales table. The reason we want to make such matches is because the information on the merchant fee charged is only available in the sales table, not in the billing table. As we will show below, the merchant fee contributes a significant amount to the overall rate of return that the credit card company earns on installments. However the rate of return on installments quoted above are *net* of the merchant fee. That is, these are the effective rates of interest that the customer paid for the installment loan. The company earns a much larger rate of return when we also factor in the merchant fee it earns at the time of the installment transaction.

In addition to installments, the company allows its customers to borrow on *cash advances*. We observe 11,818 such transactions in our data. These are typically of shorter duration than installments: the average duration of a cash advance is 45 days. The interest rates for such loans is also typically higher than for installments: it averages 24% compared to an average of 15% for installment transactions that are done at a positive interest rate (i.e. excluding the free installment transactions). The average amount of cash advances, $734, is more than twice as high as the average installment purchase done at a positive interest rate, $352. However this ranking is reversed in the upper tails of the distributions of purchases and cash advances: the largest cash advance in our data was $8300 whereas the largest installment purchase done at a positive interest rate was $15,740.

Because the motives for cash advances are likely to be different than for installment purchases and because cash advance terms are shorter and zero interest cash advance opportunities were not offered to the the company's customers (at least in our data for our sample of customers) we have chosen to limit our analysis to the choice of installment term and leave the analysis of customers' choice of cash advances to future work.

For each credit card purchase we have the following information: customer ID, types of credit card (regular card, gold card, platinum card, debit card, check card, and etc), NSS (number of the sales slip, the unique identifier for each transaction), the type of sale (including whether the sale is a return or reversal or cancellation), the date of sale (both the date of the actual sale and the date it was "posted" to the credit card), the merchant fee earned by the credit card company, and a code for the merchant type, which will be -1 for merchants that are not "in network" (i.e. for which the credit card company does not have a formal merchant agreement but does the transaction via a competing credit card's network and merchant agreement as discussed above). The sales data also include the installment term chosen if the purchase was an installment sales transaction, and the up-front cash advance fees in case of cash advance transactions. Overall, we have a total of 182,742 observations for 884 customers. The average number of transactions per customer is therefore approximately 206. Figures 4, 5 and 6 below present the distribution of the transaction amounts or ordinary (non-installment) sales, installment purchases done at a zero interest rate, and installment purchases done at a positive interest rate.

We see that, as expected, the average installment purchases are sigificantly larger than the average non-installment purchase: on average interest-free installments are four times larger and positive interest installments are seven times larger than ordinary credit card purchases. However already we can see the *free installment puzzle* in figures 7 and 8: the average size of a positive interest rate installment is more than 75% larger than the average installment done under a zero interest rate. Economic intuition would suggest that installments done at a lower interest rate, and particularly at a *zero* interest rate should be significantly *larger* than those done at a positive interest rate.

Figure 7 plots the cumulative distribution of non-installment purchases, as well as zero and positive-interest installments. We see a striking pattern: the distribution of positive-interest installments *stochastically dominates* the distribution of zero-interest installments, and this in turn stochastically dominates the distribution of non-installment purchases. Again the latter is to be expected: we would expect consumers to put mainly their larger expenditures on installment and the remaining smaller charges as regular, non-installment credit card charges. However the surprising result is that installments done at a positive rate of interest are substantially larger than installments done at a zero interest rate, at *every quantile* of the

Figure 4: Distribution of non-installment credit card purchases

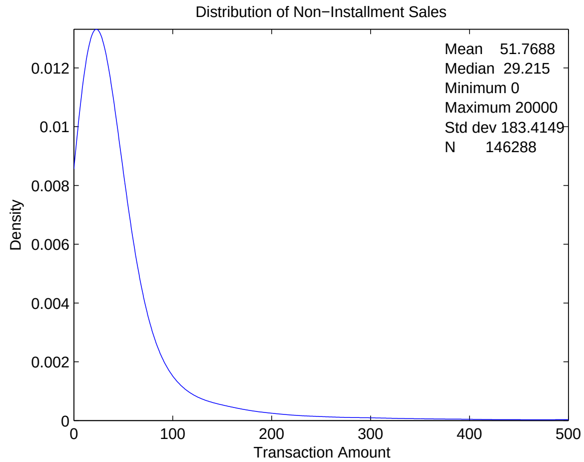


Figure 5: Distribution of positive interest installment purchases

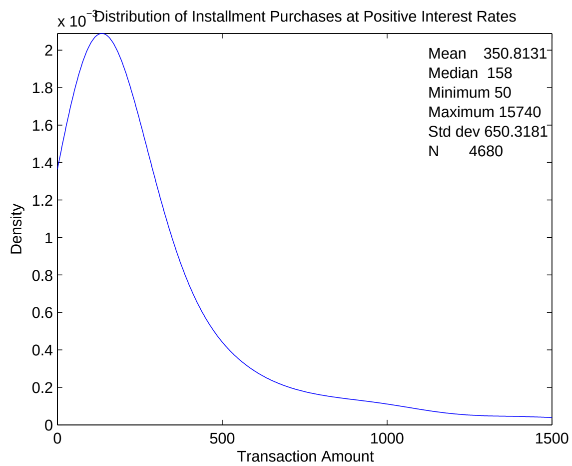


Figure 6: Distribution of zero interest installment purchases

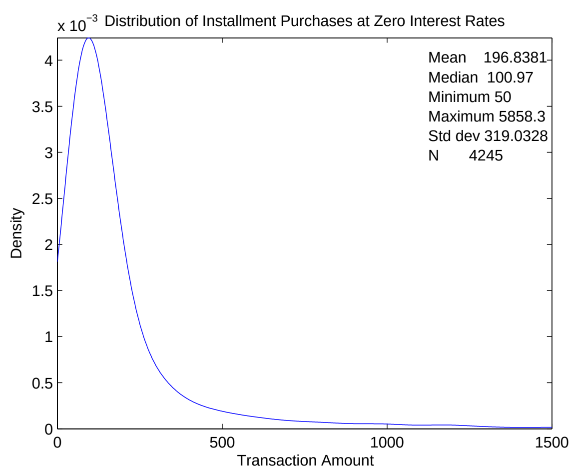
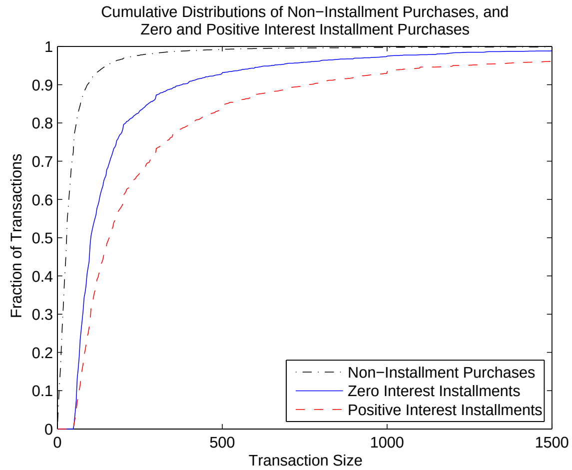


Figure 7: Cumulative Distributions of Credit Card Transaction Amounts



respective distributions. For example, the median installment at positive interest rates is nearly 60% larger than the median installment done at a zero interest rate.

In summary, the vast majority of transactions in our sales dataset, 87%, are regular (non-installment) credit card purchase transactions. These tend to be smaller in size with an average size of $50. The remaining transactions consist of cash advances (7% of the transactions) and installments (6% of the transactions). The installments we observe are roughly equally divided between zero interest and positive interest transactions. Specifically, for the subset of installment transactions that we are able to match to the billing table (which enables us to determine the interest rates actually paid, which are not contained in the sales table), approximately 47% of the installments are at zero interest and the remaining ones are done at a positive rate of interest.

Figures 8 and 9 show the distribution of internal rates of return that the credit card company earns on these installment sales, before and after accounting for the merchant fee. Recall that the internal rate of return is the (continuous time) rate of interest that sets the net present value of the cash flow stream associated with an installment purchase to zero. The credit card company experiences a cash outflow (to the merchant for the amount of the purchase) on the date the customer makes the purchase which we normalize as “day 0”. At the same time the firm received a cash inflow equal to the merchant fee received, which is actually an amount discounted from the amount paid to the merchant (if the merchant is not in-network, then the discounted payment is made to the credit card company that handles the transaction). Then at the next n statement dates the credit card company receives cash inflows equal to the repayment

Figure 8: Distribution of Rates of Return on Installments, Net of Merchant Fee

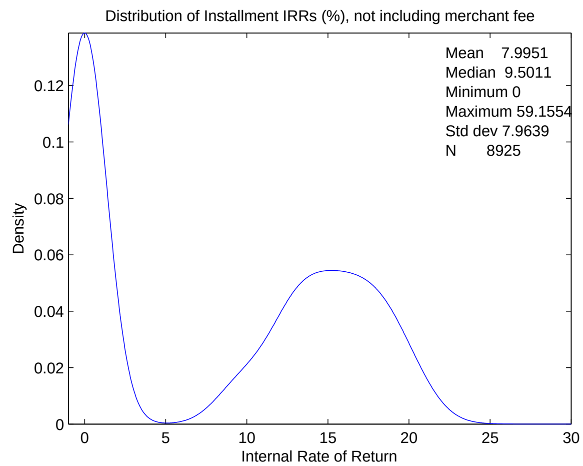
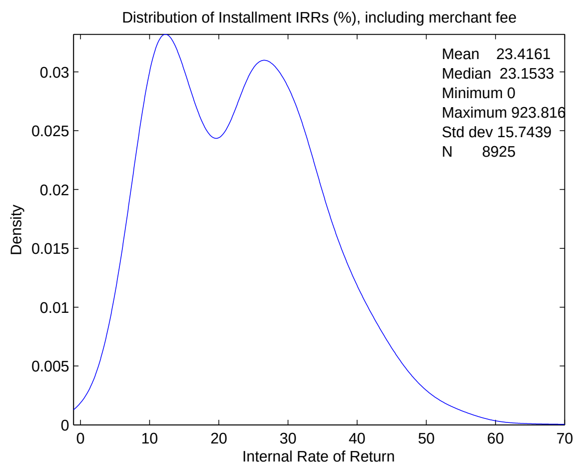


Figure 9: Distribution of Rates of Return on Installments, Including Merchant Fee



of “principal” plus interest on the installment loan.

Figure 8 shows the distribution of internal rates of returns when the merchant fee is not accounted for. This distribution is effectively the distribution of interest rates charged to the company’s customers. We see the pronounced bi-modal distribution reflecting the fact that roughly 50% of installment purchases are done at a zero percent interest rate and the other half is done at a positive interest rate. As noted above, the mean interest rate for positive interest rate installments is 15.25%.

However figure 9 shows that when we add the merchant fee, which provides the distribution of gross returns that the credit card company earns on its installment loans, we see the distribution of returns is shifted significantly to the right. Even with the “free installments” included, the company is earning an average rate of return of 23% on its installment loans, and for the positive interest installment loans the average internal return inclusive of the merchant fee is 31.4%. Of course, these calculations do not include *defaults*. However fortunately for the credit card company we studied, there were only 23 individuals out of the 938 in our sample who defaulted and whose credit card accounts were sent to collection. We cannot determine the amount of the unpaid balances that the company was ultimately able to recover from these 23 individuals, however even if all 23 were declared complete losses, including the losses into the distributions in figures 8 and 9 would not significantly diminish the estimated rates of returns that the company earns on its installment loans. Overall, we conclude that at least for this company, the installment loan business is a very good one: it pays very high rates of return with relatively low risk of default.

Already, our analysis of the credit card data in this section leads to a number of key conclusions. First, we already see the “free installment puzzle” emerging by comparing the distributions of expenditures for zero interest installments to the corresponding distribution of positive interest installments. We showed that the latter distribution stochastically dominates the former distribution, so that at every quantile in the distribution, these customers are spending more on installments that come with a large interest rate than for installments that are offered at an interest rate of zero. Secondly, we showed that the company is highly profitable and that merchant fees contribute in an important way to the overall profitability of the firm.

In fact, when we computed the (undiscounted) revenues of the firm for the 938 customers we analyzed, we found that merchant fees amounted to 36% of the total revenues received from these customers. It seems likely that the company sees merchant fees as a major component of its profits, and due to the structure of payments in this country, it places great importance on rapid growth, both in absolute and in terms of its market share, as the key to its future success.

This in turn creates strong incentive for credit card companies to try to attract new customers and to stimulate the credit card spending of its existing customers by offering free installment opportunities. However this only heightens the basic puzzle: if consumers appear to be spending *less* per transaction on the free installment opportunities they are offered in comparison to their average transaction sizes when they pay the full interest rate, what evidence is there that free installments are really stimulating spending or enabling the company to attract a significant number of new customers?

Before we go into a more focused empirical analysis directed at the specific issue of attempting to estimate the "demand for credit" we find it useful to present some additional distributions and scatterplots that reveal some additional important facts and features that our data that our empirical models will need to explain. In particular, we present some further data that helps us to understand which types of individuals are the most likely users of installment credit.

Figures 10 and 11 show the distribution of the number of credit card *transactions* and the *share of all credit card spending* done as installment purchases. We see that while installments are less than 9% of all credit card transactions, they account for more than 25% of all credit card spending.

Of course, this is due to the fact that the average credit card purchase is $74 while the average installment purchase is $364, with the full distributions of the average purchase and installment transaction sizes over the consumers in our sample plotted in figures 12 and 13. Thus, consumers generally pay for much larger items (or more expensive baskets) on installment, but choose to pay smaller amounts in full at the next statement date. We are also struck by the much greater skewness of the distribution of installment purchases relative to that of credit card purchases as a whole.

Our analysis reveals a substantial degree of heterogeneity across credit card customers in their propensity to make use of installments to pay for their credit card purchases. Overall our analysis suggests that the best single measure of the propensity to use installments is not the mean fraction of transactions done via installment, but rather the mean share of credit card purchases paid for by installment. Hereafter we will refer to the latter measure as the *installment share.* Now we will turn to a series of scatterplots that relate the installment share to other covariates we observe in our credit card data set.

Figures 14, 15 and 16 present scatterplots (with the central tendency of the data indicated by a local linear regression fit to the data) of how the installment share relates to various measures of creditworthiness. Figure 14 plots the installment share against customer credit scores, using the company's internal (proprietary) credit scoring system where a score of 1 represents the best possible creditworthiness and 12

Figure 10: Distribution of the Fraction of Credit Card Transactions done as Installments

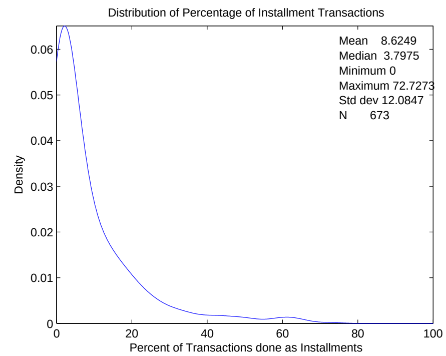


Figure 11: Distribution of the Share of all Credit Card Spending done as Installments

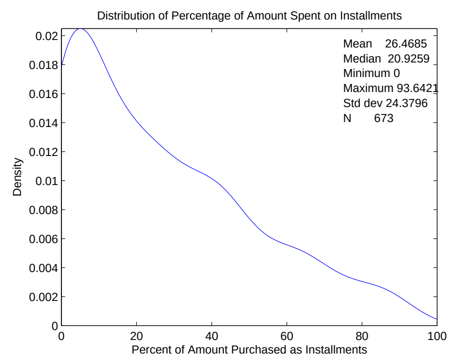


Figure 12: Distribution of the Average Amount of a Credit Card Purchase across Customers

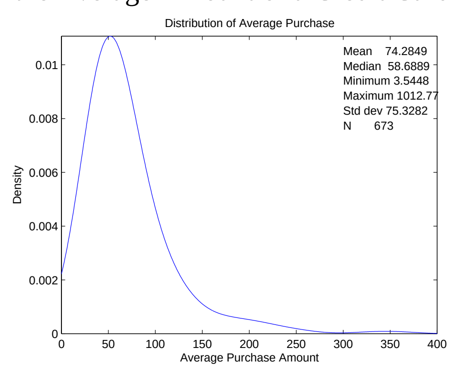
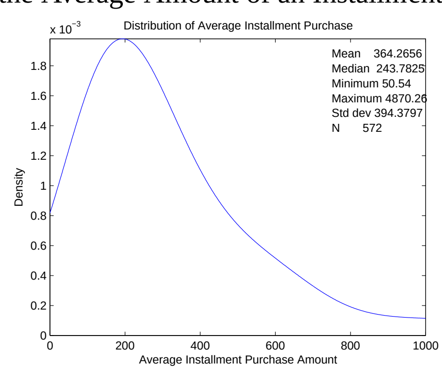


Figure 13: Distribution of the Average Amount of an Installment Purchase across Customers



is the worst. Customers who have credit scores in this range are still allowed to borrow on installment and face no credit limits. However consumers who are in the process of collection will have their credit card borrowing and spending privileges suspended and they show up in our data set as having a credit score of 0. We see generally negative correlation between the credit score and the installment share (remember that higher credit scores indicate worse credit, so the relationship in figure 14 is actually positively sloped).

We see figure 14 as a potential first indication of possible credit constraints, or at least *high demand for credit* among the customers that are heavy installment spenders. Perhaps their poor credit score indicates that they are also regarded as poor credit risks to other lenders, and as a result of this, they are forced to make heavier use of the installment credit facility of this credit card company at relatively high rates. On the other hand, the customers with the best credit scores also generally the least heavy users of installment, which could be an indication that they are not liquidity constrained, or have other lower cost sources of access to credit elsewhere.

Figures 15 and 16 illustrate the incidence of late payments. Figure 15 shows that the average number of late payments per customer is positively correlated with the installment share, and figure 16 shows that the number of *seriously late* payments (i.e. payments that are 90 or more days past due, or at about the threshold where the company suspends credit card charging privileges) is also positively correlated with the installment share. These figures confirm the conclusion we obtained in figure 15, namely, that customers who are heavy users of installment spending are also worse credit risks.

Figures 17 and 18 relate the installment share to three separate indicators of the type of installment spending that customers do. Figure 17 presents a scatterplot of the ratio of the size of a typical installment purchase to the typical credit card purchase. As we noted previously, credit card customers generally pay for only relatively large purchases on installment, and pay for the smaller transactions in full at the next statement date. We see that as a function of the installment share, the low intensity installment users tend to buy items on installment that are between 4 and 6 times as large at their typical credit card purchase. However for the heaviest users of installment spending this ratio falls to less than 3, which potentially indicates a more “desperate” individuals who are more likely to pay for smaller “everyday” items by installment.

Figure 18 shows a scatterplot of the ratio of the installment balance to the average statement balance as a function of the installment share. Of course, this ratio is positively correlated with the installment share is almost definitional, but the figure does show that the heaviest installment users carry installment balances

Figure 14: Customer-Specific Average Credit Scores by Installment Share

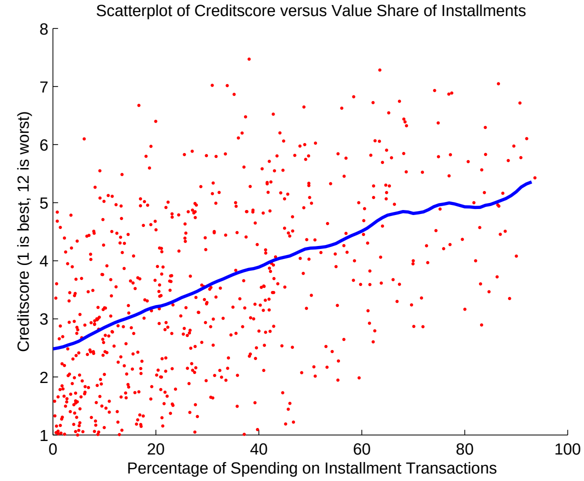


Figure 15: Number of Late Payments by Installment Share

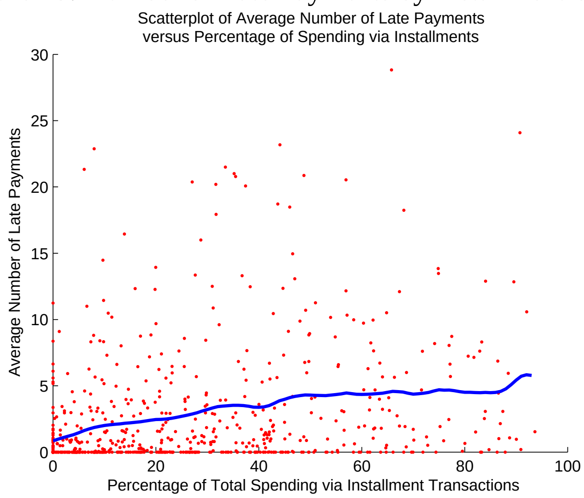


Figure 16: Number of Seriously Late Payments (over 90 days) by Installment Share

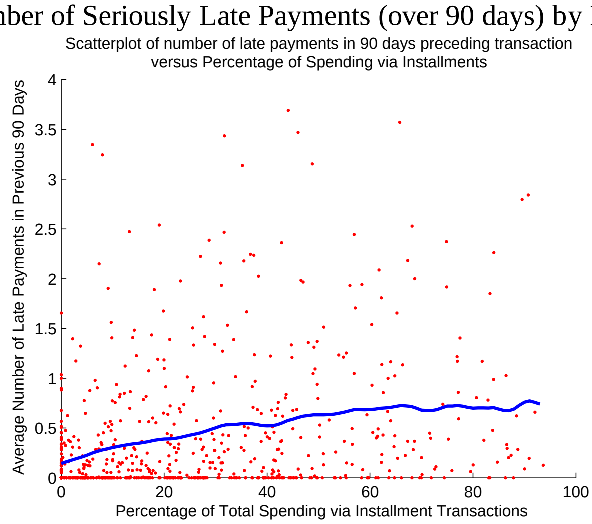


Figure 17: Ratio of Installment Size to Typical Purchase Size by Installment Share

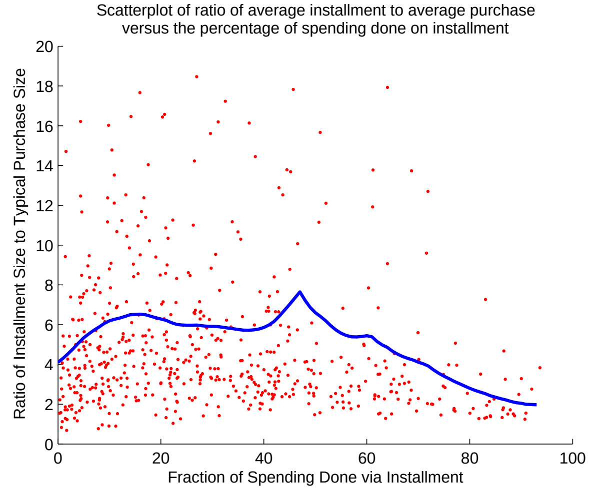


Figure 18: Ratio of Installment Balance to Average Statement Balance by Installment Share

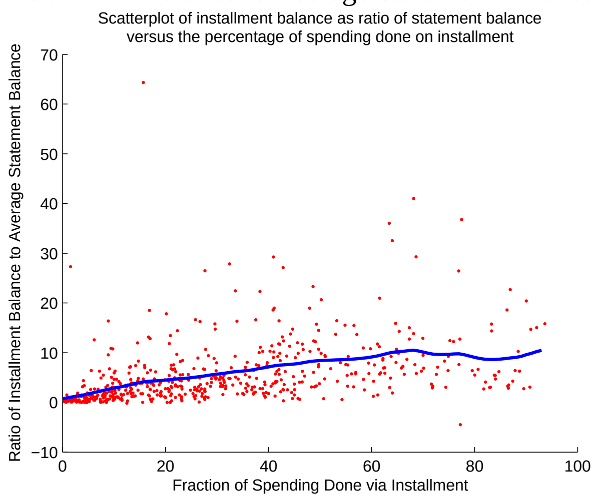
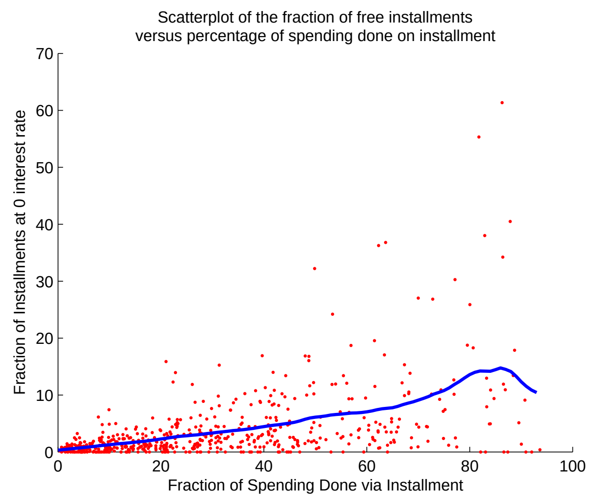


Figure 19: Fraction of Installment Transactions done as Free Installments by Installment Share



that are on average 10 times larger than their typical monthly credit card balances (statement amounts).

Figures 19 and 20 relate the usage of free installments to the installment share. In figure 19 we see that the fraction of installment transactions done as free installments is positively correlated with the installment share. The previous figures in this section lead to an impression that the heavy installment spenders are relatively desperate for credit, and thus, it would seem logical that they are the ones who would be most likely to take the greatest advantage of free installment opportunities when they encounter them. The upward sloping relationship in figure 19 is consistent with this interpretation, and shows that the heaviest installment users are doing as much as 20% of their installment purchase transactions as free installments (i.e. at 0% interest rate).

Figure 20 shows a similar relationship but instead of plotting the fraction of installment transactions that are done as free installments it shows the share of installment spending that is done via free installments. Both of these graphs show a similar pattern, namely that the customers with the highest installment shares are doing about 15-20% of all of their installment transactions and 15-20% of all installment spending via free installment offers.

We conclude this section with figures 21 and 22 that give us some insight into the profitability of the "free installment marketing strategy" used by this firm. Recall from section 2 that we suggested that the company's use of free installment offers seems motivated by a desire to increase its customers' use of its credit cards in an attempt to increase its credit card market share, since doing this increases its leverage in setting merchant fees, which we showed in section 3 are a major component of the high profitability of this company. However we have also shown in this section that the customers that are most likely to act on the free installment offers are those with worse credit scores and higher incidence of late payments. As such, the use of free installments as a promotional device may have the perverse effect of offering free credit to the company's least creditworthy customers, and this group may be the most likely to default. This creates the possibility that free installments might be a relatively ineffective and/or highly costly means of increasing credit card usage.

Figure 21 plots the average internal rate of return on all installment transactions (including free installments) against the installment share. We see that this curve is upward sloping, which indicates that even though the "installment addicts" are the ones most likely to be taking up the free installment opportunities, the interest rates that they pay on their positive interest installment transactions are rising sufficiently fast with the installment share that it counteracts the "free installment effect" so that overall average install-

Figure 20: Share of Installment Spending Done as Free Installments

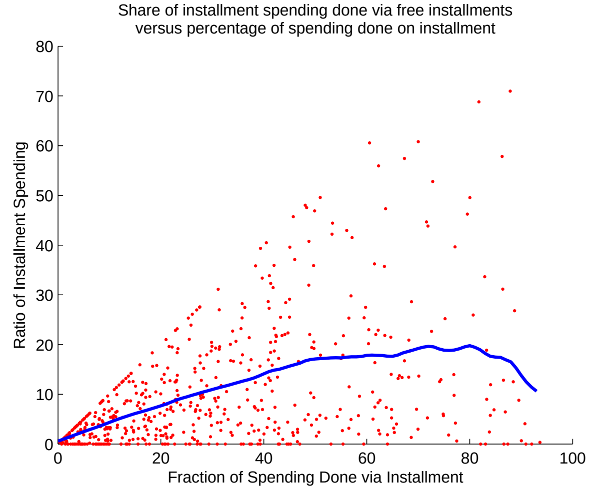


Figure 21: Average Internal Rates of Return on Installments by Installment Share

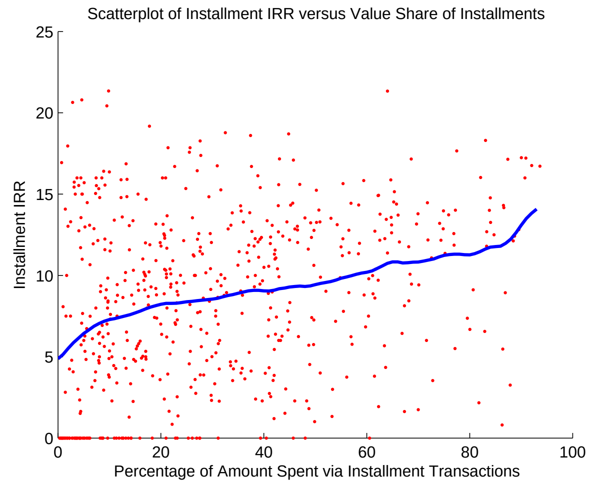
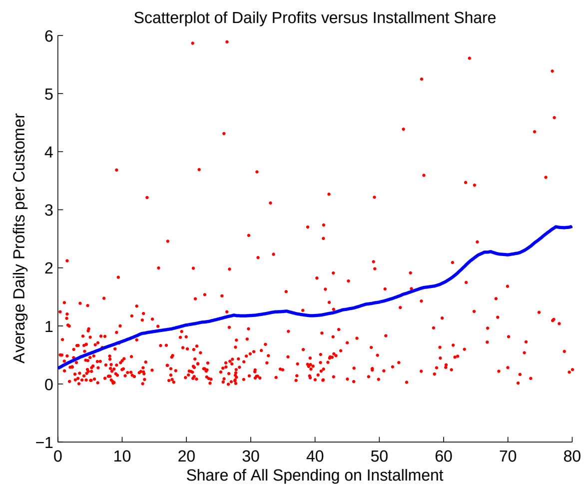


Figure 22: Customer-level Daily Profits by Installment Share



ment interest rates paid by its customers increase monotonically as a function of the installment share. Of course the reason for this is likely to be related to the fact that the customers with high installment shares have significantly worse credit scores, and as we will show in section 4, the interest rates that customers pay is a monotonically increasing function of their credit score (i.e. customers with higher scores, which indicate worse credit risks, pay higher interest rates).

Figure 22 plots the average daily profits for each consumer against the installment share. This figure indicates a pronounced upward sloping relationship between the installment share and the profitability of customers. If we believe this is the relevant figure to focus on, then the company's free installment marketing policy seems rational and well targeted: it appears to be succeeding in having the biggest impact on the most profitable customers, but these customers also happen to have worse credit scores and present higher credit risks.

However given the relatively small number of observations and the relatively large number of outliers, we think it is hazardous to come to any definite conclusion one way or the other about the wisdom of free installments at this point. As we noted in the previous section, we cannot address with our data a crucial missing piece of information that would be needed to provide a fuller answer to this question: to what extent does the knowledge of free installments cause customers to increase their spending? Recall that we are doing our analysis *conditional* on the decision to purchase a given item. We would need additional information to determine whether the existence and knowledge of free installment opportunities causes the company's customers to go to stores more often, purchase more at a given store than they otherwise would, or increase their likelihood of using the company's credit instead of paying for the item using a competing credit card or cash.

3 Reduced-Form Approaches to Inferring the Demand for Credit

The data we have would appear ideal for empirically modeling *the conditional demand for credit* — at least as it pertains to relatively smaller scale short term borrowing decisions. As we noted above, we define the conditional demand for credit as the demand to finance a given credit card purchase through borrowing rather than to pay the amount purchased in full at the next purchase date. It is conditional on having made a decision to make a given purchase of a given size in the first place. As we noted above, we do not have the appropriate data that would enable us to model how access to borrowing and how the interest rate schedule that a customer can borrow at also affects the frequency and amounts of purchases. We would need additional sources of data, then, to attempt to estimate the fuller *unconditional demand for credit.*

To make this a bit more precise, we introduce a bit of notation. Let c denote the decision by the consumer to pay using the company's credit card (as opposed to paying by cash, or using some other credit card). Let r be the interest rate charged to a customer with observed characteristics x for purchasing via installment credit. As we show in more detail below, we should interpret r as an entire *interest rate schedule* since the customer can ordinarily choose the term of the installment loan and thus faces a consumer-specific "term structure" of interest rates. Consider the demand for credit via the company's credit card c over a specific interval of time, say one month. The (unconditional) expected demand for credit by a single customer with characterics x, $ED(r,x,c)$ (where x includes variables such as the customer's credit score, spending history, and might also include information on interest rates offered by competing credit cards or interest rates for other sources of credit) can be written as follows

$$ED(r,x,c) = \left[\int_0^\infty a[1 - P(1|a,r,x,c)] f(a|x,r,c)da\right] \pi(c|r,x)EN(x,r). \tag{1}$$

where $P(1|a,r,x,c)$ is the probability that a customer will choose to pay for a purchase amount a in full at the next statement date given the interest schedule r, the consumer characteristics x and the decision to use the company's credit card c to carry out the transaction. We let $\pi(c|r,x)$ denote the customer's decision to use the credit card company c's credit card to pay for the transaction, and $f(a|x,r,c)$ denotes the density of the amount purchased using the company's credit card during any given shopping trip. Finally $EN(x,r)$ denotes the expected number of shopping trips that the customer makes during the specified interval of time. The overall expected demand for credit from the customers of credit card company c is then just the sum over the customer-specific expected demand curves $ED(r,x,c)$.

The data we have are not sufficient to estimate the objects $\pi(c|r,x)$ or $EN(x,r)$. Separate survey data would have to be collected that would enable us to study the purchase habits of a sample of the company's customers, and how something like free installment offers during a given period of time might affect the number of shopping trips they make (thus enabling us to estimate $EN(x,r)$), or the likelihood that they will use the company's credit card c to pay for the purchase (thus enabling us to estimate $\pi(c|a,r,x)$).

However since we do observe all of the purchase amounts that a given consumer makes during any given shopping trip where the customer uses the company's credit card, we can potentially estimate $f(a|x,r,c)$. Further, since we also observe customers' choices of whether to purchase on installment or whether to pay the amount a in full at the next statement date conditional on having decided to use the company's credit card, we can potentially estimate the *installment choice probability* $P(d|a,r,x,c)$, where the option $d = 1$ indicates a choice to pay the purchase amount a in full at the next statement date. If so, then by segregating customers' purchases into those that are paid in full at the next statement date and those that are paid on installment, we can estimate two conditional densities, $f_0(a|x,r,c)$ (i.e. the distribution of purchase amounts that are paid in full at the next statement date) and $f_1(a|x,r,c)$ (the distribution of purchase amounts that are paid for by installment). We have already presented the unconditional analogs of f_0 and f_1 in figures 12 and 13 of section 2, where we showed in particular that the average size of an installment purchase was nearly 5 times larger than the average size of a non-installment transaction. Since f_0 and f_1 are conditional distributions, we can write them according to the usual formulas of probability theory

$$\begin{aligned} f_0(a|x,r,c) &= \frac{P(1|a,r,x,c)f(a|x,r,c)}{\int_0^\infty P(1|a,r,x,c)f(a|x,r,c)da} \\ f_1(a|x,r,c) &= \frac{[1-P(1|a,r,x,c)]f(a|x,r,c)}{\int_0^\infty [1-P(1|a,r,x,c)]f(a|x,r,c)da}. \end{aligned} \tag{2}$$

Thus, we can at least use our data to estimate the *conditional expected demand for credit* $ED_1(r,x,c)$ which we define as

$$ED_1(r,x,c) = \int_0^\infty a f_1(a|x,r,c)da. \tag{3}$$

Just as we expect the unconditional demand curve to be a downward sloping function of r, we also expect the conditional demand for credit to be downward sloping in r because we expect customers to borrow larger amounts on installment when the interest rate is lower. Even if the distribution of purchase sizes was unaffected by r (i.e. if $f(a|x,r,c)$ was not a function of r), a downward sloping demand would still follow if the probability that a customer chooses to pay the purchase amount a in full at the next statement date is an increasing function of r (in which case the customer's credit demand is nothing beyond that inherent in the typical "float" i.e. the lag between buying an item with a credit card and paying for it at the next statement date).

It follows that if we restrict attention to the subset of transactions that a customer purchases on installment credit, we have the regression equation

$$\tilde{a}_i = ED_1(r,x,c) + \tilde{\varepsilon}_i \tag{4}$$

where $\tilde{a}_i$ is the amount borrowed in the i^{th} installment transaction made by the customer, and $\tilde{\varepsilon}_i$ is a residual satisfying $E\{\tilde{\varepsilon}_i|r,x,c\} = 0$. We refer to the regression equation (4) as the conditional demand curve for credit, and it seems like a natural place to start is to estimate this regression by ordinary least squares. However rather than attempt to specify parametric functional forms for the underlying components of the regression function $ED_1(r,x,c)$, i.e. the probability $P(1|a,r,x,c)$ and the density $f(a|x,r,c)$ which would result in a specification that is nonlinear in the underlying parameters, it is also natural to start by estimating a flexible linear-in-parameters approximation to the regression function $ED_1(r,x,c)$.

However, perhaps not surprisingly, we find that when we do these ordinary least squares regressions for every specification we tried where the dependent variable is the amount of an installment purchase and for different combinations of right hand side (x,r) variables, we always found that the regression predicted a strong, and statistically significant *positive relationship* between the expected amount of installment borrowing and the interest rate r. to have a *positive and statistically significant coefficient.* That is, the regressions are suggesting that the *conditional (expected) demand for credit is upward sloping!*

Of course, the ordinary least squares regression results are likely to be spurious due to the *endogeneity of the interest rate.* That is, we can imagine that there are *unobserved characteristics* of consumers that affect both their willingness/desire to make purchases on credit and the interest rate they are charged. In particular, we would imagine that customers who are *liquidity constrained* and who might exhibit *bad characteristics* that can lead them to simultaneously wish to borrow more but at the same time constitute a *higher credit risk* will have worse credit score and therefore face a higher rate of interest, but will still have a higher propensity to borrow due to their liquidity constraints and a dearth of alternative, better borrowing options. Indeed, as we already showed in figure 14 of section 2 that there is a strong correlation between the fraction of spending on installment credit and the credit score: individuals with worse credit scores tend to do a higher fraction of their credit card purchases on installment. Given the monotonic relationship between credit scores and installment interest rates, it is not hard to see why the regression estimate of the installment interest rate is positive and statistically significant.

We attempted to deal with the endogeneity problem using the standard arsenal of “reduced form” econometric techniques, including *instrumental variables.* In particular, we have access to daily interest rates that measure the “cost of credit” to the bank for the loans it makes to its customers, including 1) *the certificate of deposit CD rate* and 2) *the call rate.* The latter is an interbank lending rate for “one day loans.” Both the CD rate and the call rate change on a daily basis. We use these rates as instrumental variables on the theory that in a competitive banking market, no single bank can affect the CD or call rates, and thus changes in these rates can be regarded as exogenous changes in the cost of credit that the credit card companies ultimately "pass on" to their credit card customers. However the instrumental variables (two stage least squares) estimate of the coefficient of the interest rates the company charges its customers becomes *statistically insignificant* as you can see in table 1 below. The coefficient estimates of the interest rate r are highly sensitive to whether we include all installment transactions (including those with $r = 0$ or just those with $r > 0$. We obtain a highly negative but statistically insignificant point estimate in the former case, and positive and statistically insignificant estimate in the latter.

We define the average treatment effect (ATE) as our "parameter of interest" even though our actual interest is to estimate the conditional demand curve for credit. Given the poor results from instrumental variables estimation, we are now willing to settle for a much less ambitious goal: can we even show that people will borrow more when offered 0% interest compared to when they must pay high positive rates of interest? The ATE is simply an estimate of the difference between mean borrowing for the treatment group who were offered zero interest

We do not really believe the inferences from our instrumental variables regressions, or the suggestion that we have a unique finding that the demand for credit is some sort of *Giffen good.* After all, if the firm believed that charging higher interest rates causes its customers to spend *more,* why would it offer free installment opportunities? Instead we believe that the reduced-form results are spurious, and in particular both the CD and call rate are *weak instruments.* Indeed, not only are they weakly correlated wtih consumer interest rates, we find that the CD and call rates are *negatively correlated* with the interest rates the firm charges to its customers. We view this as evidence that the credit card market is not "competitive" and the are substantial "markups" in the interest rates charged to customers over the cost of credit to the banks, and this markup is driven more by customer specific risk factors and by competitive trends within the credit card market itself than by the the much smaller day to day fluctuations in the CD and call rate. The latter have hovered in a fairly narrow band between 3 or 4 percent over the period of our analysis whereas installment interest rates vary much more widely across customers and over time as their credit scores change, ranging from as low as 5% to 25% or higher.

The next approach we considered in order to try to infer the "causal effect" of interest rates on the demand for credit was *matching estimators.* The idea behind these estimators is to compare the average amount purchased by individuals who were offered free installments (the "treatment group") with a corresponding and "similar" set of individuals who took out installment loans when purchasing from similar

Table 1: Instrumental Variables-Fixed Effects Regressions of Conditional Demand for Credit
Dependent variable: $log(a)$ where a =amount borrowed. Amounts in parentheses are
P-values for tests of the hypothesis that the coefficient/statistic is zero.

Item	Specification 1	Specification 2	Specification 3	Specification 4
Instruments	CD rate	CD rate	CD rate	CD rate
	credit score	credit score		
Fixed effects	yes	yes	yes	yes
Free Installments	yes	no	yes	no
Variable	Estimate	Estimate	Estimate	Estimate
r	0.965	-72.903	0.739	-102.20
	(0.249)	(0.591)	(0.382)	(0.628)
credit score	0.001	-0.002		
	(0.442)	(0.835)		
$d=2$	0.314	2.733	0.317	3.695
	(0.000)	(0.531)	(0.000)	(0.588)
$d=3$	0.896	4.9644	0.912	6.543
	(0.000)	(0.690)	(0.000)	(0.561)
$d=4$	1.028	5.434	1.042	7.099
	(0.000)	(0.474)	(0.000)	(0.549)
$d=5$	1.06	6.623	1.061	8.668
	(0.000)	(0.472)	(0.000)	(0.547)
$d=6$	1.828	7.172	1.840	9.243
	(0.000)	(0.450)	(0.000)	(0.533)
constant	11.500	20.559	11.519	24.104
	(0.000)	(0.216)	(0.000)	(0.349)
σ_u	0.651	1.276	0.652	1.708
σ_ε	0.656	1.788	0.657	2.420
ρ	0.495	0.337	0.496	0.331
Sample size	8183	4109	8078	4049
F-test ($u_i = 0$)	$F(613) = 8.03$	$F(474) = 0.97$	$F(598) = 8.08$	$F(464) = 0.53$
	(0.00)	(0.687)	(0.00)	(1.00)
Hausman test	$H(8) = 6.54$	$H(8) = 1.96$	$H(6) = 4.45$	$H(6) = 0.23$
	(0.59)	(0.96)	(0.61)	(0.99)

merchants at similar periods of time but at a positive interest rate (the "control group"). Since there are many individuals in our sample for which we observe a large number of installment transactions (these are the heavy installment "addicts" that we discussed in the previous section who have installment shares in excess of 50%), we can even use a number of individuals as "self-controls" — that is we can compare the average size of free installments with the average size of installments done at positive interest rates for the same individual, where we do additional matching by selecting a set of free installments and positive interest rate installments that were done at approximately the same intervals of time and from approximately the same set of merchants.

Specifically, we focuses on attempting to estimate the "average treatment effect " (ATE) where the "treatment" in question is offering a customer a free installment borrowing opportunity, which we denote as $r = 0$. The ATE is defined as the difference in the expected borrowing between the treatment group $r = 0$ and control group $r > 0$

$$\text{ATE} = E\{a|r = 0\} - E\{a|r > 0\}, \tag{5}$$

where a is the amount borrowed and r is the interest rate. The idea behind the matching estimator is that if we are able to match a sufficiently large number of customers in the treatment and control groups on a sufficiently narrow set of criteria X such that we can plausibly assume that the "assignment" of the "treatment" $r = 0$ is essentially random for the matched individuals/transactions, then we can infer what the installment spending for a treated person would be by taking the mean installment spending for the matched individuals in the control group (and vice versa) and essentially estimate the ATE as if it were a result of a classical controlled randomized experiment for subsets of matched individuals and transactions and averaging these match-specific treatment effects across all matched groups in the sample. The validity of this approaches depends on a conditional independence assumption known by the (unfortunate) name, "the unconfoundedness assumption" (or also, the "strong ignorability assumption"). The table below presents our estimates of the ATE, which we would expect to be positive if the demand for credit were downward sloping.

We can see from table 2 that regardless of how we do the matching of individuals/transactions the estimated treatment effects are all estimated to be of the *wrong sign and highly statistically significant.* The estimated treatment effects become increasingly negative as we use increasingly relaxed criteria for matching individuals, but overall given the magnitude of the estimated standard errors for the estimated ATE's, there is no strong evidence that the various estimates are statistically significantly different from

Table 2: Effect of Free Installments: Results from Matching Estimators

Matching Criteria	Estimated ATE	Standard Error	P-value for H_o : ATE $= 0$
customer, credit score CD rate, merchant code	-\$56.60	\$15.20	0.000
customer, credit score merchant code	-\$69.51	\$16.45	0.000
customer, merchant code	-\$79.33	\$19.93	0.000
customer	-\$76.72	\$18.75	0.000
merchant code	-\$61.07	\$16.00	0.000

each other. However we can strongly reject the hypothesis that the ATE is zero. Thus, we are left with the paradox that the matching estimator predicts that free installment opportunities cause customers to *reduce the amount of their borrowing* and therefore, the matching estimators imply an *upward sloping demand for credit.*

4 Exploiting the Quasi-Random Nature of Free Installment Offers

In view of the failure of the various reduced form methods that we tried in the previous section we started to think "outside the box" for other ways to provide more credible and econometrically valid estimates of the conditional demand for credit. Our goal was to develop an approach was that is capable of exploiting the information contained in the company's use of free installment offers as a *quasi random experiment.*

Note that we already tried to do this, albeit unsuccessfully, in the previous section, where we applied one of the standard approaches in the "treatment effects" literature, namely the use of matching methods. Unfortunately the matching estimators were all strongly statistically and economically significant but of the wrong sign. Although the quasi-random nature of the way the credit card company offers free installment offers to its customers does provide a strong degree of *prima facie* plausibility for the validity of the key conditional independence assumption that justifies the use of matching estimators, the fact that there is a great deal of *self-selection* in which individuals choose to take free installment offers suggests that there could be an important problem of *selection on unobservables* that could invalidate the conditional independence assumption and cause the matching estimators to result in spurious estimates. We now present an approach that can exploit the quasi random nature of free installment offers that is also robust to the possibility of selection on unobservables. Unfortunately, in the absence of further data, or without the ability to conduct randomized, controlled experiments, our ability to exploit free installments as a quasi random does require some degree of modeling and assumptions.

Consider first what would be possible if we had data from a *randomized controlled experiment* (RCE). Though the company we are studying has not done this to our knowledge, one could imagine that the company could be convinced to undertake such a study to get better estimates its customers' demand for installment credit. For example the recent study (Alan et al. [2011] is an example where an enlightened credit card company did choose to undertake a large scale RCE to better understand its customers' demand for credit. In a classical RCE the company would randomly assign a subset if its customers to a control group and a treatment group. Individuals in the control group would continue to receive the same interest rates for installments that they receive under the *status quo* while individuals in the treatment group would be offered randomly assigned alternative installment interest rates. The alternative interest rates could be either higher or lower, or even zero, and by comparing the demand for installment loans for the treatment and control groups, we could essentially use the random assignment as a valid "instrument" to help solve the problem of endogeneity in the interest rate, and make valid inferences about the conditional demand for credit.[2]

In order to exploit the free installment promotions the credit card offers as a type of *quasi random experiment* (QRE) we can no longer do simple comparisons of responses (e.g. demand for credit) of "control" and "treatment" groups. In particular, in our data while we can be sure that individuals who accepted free installments were offered the "treatment", we cannot simply assume that individuals who did not choose free installments are are in the "control group" (i.e. were not offered free installments) since some of these individuals might have been offered free installment opportunities, but decided not to accept them. Therefore, in order to fully exploit the information provided by the existence of free installment offers, we do have to undertake some additional modeling and make some additional assumptions.

In particular, the self-selected nature of customers' decisions to take advantage of free installment offers is compounded by another potentially serious measurement issue, namely *censoring.* That is, *our data only allows us to observe free installment offers when customers actually choose them, however for all other non-free installment transactions, we cannot observe whether the customer was not offered a*

[2]Note that Ausubel and Shui [2005] analyzed data from a randomized experiment, but it was not a RCE since there were no "controls" corresponding to the subjects who were offered the "treatments" (i.e. the six introductory offers). However to a certain extent the individuals who were offered different introductory offers could be regarded as controls. For example the individuals who were offered a 7.9% 12 month introductory offer could serve as controls for the individuals who were offered the 4.9% 6 month introductory offer, but doing this only allows us to test how customers respond to one of these offers relative to the other one. They cannot tell us how the customers who accepted either of these introductory offers behaved relative to customers who were not offered either introductory offer: the company would have have to have included an explicit control group to do this —- i.e. a 7th group of customers who decided to sign up for the credit card without being offered any special introductory offer.

free installment opportunity, or if the customer was offered a free installment opportunity but the customer chose not to take it. Since we are willing to make some reasonable assumptions and put some additional structure on the credit choice problem, we can provide econometric solutions to the censoring and self-selection problems, enabling us to infer how interest rates affect the choice of installment term and the conditional demand for credit.

4.1 The Discrete Choice Model

Assume that a customer with characteristics x evaluates each transaction in terms of the *net utility* of postponing the payment of the purchase over a term of d months. The customer faces an interest rate $r(x,d)$ for borrowing over a term of d months, except that $r(x,1) = 0$, i.e. all customers get an "interest free loan" if they choose to pay the purchase amount a in full on the next statement date. We normalize the net utility of this "pay in full" option, $d = 1$, to 0. However for the installment purchase options $d = 2,3,\ldots,12$ we assume that the net utility is of the form $v(a,x,r,d) = ov(a,x,d) - c(a,r,d)$ where $ov(a,x,d)$ is the *option value* to a customer with characteristics x of paying for the purchase amount a over d months rather than paying the amount in full a the next statement date (which has an option value normalized to 0 as indicated above, $ov(a,x,1) = 0$).

The function $c(a,r,d)$ is the *cost of credit* equal to the (undiscounted) interest that the customer pays for an installment loan of amount a over duration d at the interest rate r. The net utility

$$v(a,x,r,d) = ov(a,x,d) - c(a,r,d) \tag{6}$$

can therefore be regarded as capturing an elementary cost/benefit calculation that the customer makes each time he/she makes a transaction with their credit card.

We add onto each of the net utilities $v(a,x,r,d)$, $d = 1,2,\ldots,12$ an additional Type I (Gumbel) extreme value error component $\varepsilon(d)$ that represent the effect of "other idiosyncratic factors" that affect an individual's choice of installment term that are independent across successive purchase occasions, so that the overall net utility of choosing to purchase an amount a on an installment of duration d months is $v(a,x,r,d) + \sigma\varepsilon(d)$, where $\sigma > 0$ is a scale parameter that determines the relative impact of the "idiosyncratic factors" $\varepsilon(d)$ relative to the "systematic factors" affecting decisions as is captured by $v(a,x,r,d) = ov(a,x,d) - c(a,r,d)$.[3] Examples of factors affecting a person's choice that might be in the

[3]Specifically, we assume that $\varepsilon(d)$ are "standardized" Type I extreme value random variables, standardized to have scale

$\varepsilon(d)$ term is whether there is a long line at checkout (so the customer feels uncomfortable weighing the options $d = 2,\ldots,12$ relative to doing the "default" and choosing $d = 1$), or if a customer has time-varying but uncorrelated psychological uncertainty about what other bills or payments may be due at various up-coming months $d = 2,\ldots,12$.

As is well known, when we "integrate out" these unobserved components of the net utilities we obtain a multinomial logit formula for the conditional probability that a consumer will choose an installment term $d \in \{1,\ldots,12\}$. For consumers who are not offered any free installment purchase opportunity, their choice set is the full set of 12 alternatives $d \in \{1,2,\ldots,12\}$. However for a consumer who is offered a free installment opportunity to spread a purchase a over a maximum of $\delta > 1$ payments, we will test a key *dominance assumption*, namely that all customers strictly prefer a free installment opportunity of duration δ over any positive interest rate installment of *shorter* duration, $d = 2,3,\ldots,\delta-1$. The dominance assumption implies that the probability of choosing any positive interest rate alternative $d < \delta$ is zero.

We consider and test two versions of the dominance assumption. The *strong dominance assumption* is the one described above, namely that a customer who is offered any free installment offer of maximum duration δ will never choose any duration $d < \delta$ including the option of paying in full for the amount purchased at the next statement date, which is the choice of alternative $d = 1$. The strong dominance assumption emerges as a limiting outcome if $ov(a,x,d) > 0$ and $ov(a,x,d)$ is non-decreasing in d in the limit as $\sigma \downarrow 0$, since for any free installment offer we will have $c(a,r,d) = 0$ for $d \leq \delta$ where δ is the maximum allowed duration of the free installment offer. As $\sigma \downarrow 0$, the implied choice probabilities from the discrete choice model will assign probability 0 any choice $d < \delta$, though it does not rule out the possibility that a sufficiently liquidity constrained consumer could pay a positive interest rate for a installment loan of longer duration that the maximum term δ offered under the free installment option.

We will show shortly that we can strongly reject the strong dominance assumption. In particular. while the credit card does not keep records that can enable it to precisely estimate what the overall probability of free installment offers is, company employees we did speak to are quite certain that the rate is significantly higher than 2.7%. which is the fraction of transactions we observe being done under free installment offers, and would constitute an estimate of the average probability of free installment offers in our sample if the strong dominance assumption held.

Therefore we consider and test an alternative *weak dominance assumption.* Under the weak dominance

parameter equal to 1, so $\sigma\varepsilon(d)$ is then a Type I extreme value random variable with scale parameter σ.

assumption, we assume that there may be "mental accounting costs" that might deter a customer from taking an installment offer, even if it were free, but if a customer finds it optimal to incur these mental accounting costs and choose the free installment option, then these customers will always choose a loan duration d equal to the maximum loan duration δ permitted by the company under the free installment offer. After all, since there is no pre-payment penalty, if *ex post* events make it optimal for the customer to pay off the installment balance faster than over the δ months allowed under the free installment offer, the customer is always free to do so. As we noted in the introduction, it is very hard for standard economic theories to explain why an individual would pre-commit to taking the installment for any shorter term $d \in \{2,\ldots,\delta-1\}$ when there is no apparent cost to choosing the maximal allowed term δ and choosing the maximal term gives the customer the option that has the maximal *ex post* flexibility in terms of uncertain future events that may affect his/her ability to pay off their account balance.

We do not test a third variant of the dominance assumption, namely, that if a customer were to choose an installment loan of shorter duration than the maximum duration offered, $1 < d < \delta$, the customer would always choose this loan to be at a zero interest rate rather than at a positive interest rate. We cannot test this even weaker variant of the dominance assumption because the credit card company *forces* customers to choose the zero interest installment option over the positive interest installment option whenever the duration of their installment loan is less than the maximum duration offered, δ. However customers do always have the option to choose installment loans of *longer* duration than the maximum duration of the free installment offer δ (unless $\delta = 12$) and then in such cases the customer would pay a positive interest rate to choose one of the longer installment durations $d \in \{\delta+1,\ldots,12\}$. As we will see, our model allows for this possibility and predicts that it will occur, though the probability that it happens is small.

If we observed whether consumers had a free installment option *regardless of whether or not they choose the free installment option* our life would be much simpler. Then we could write a *full information likelihood function* that is the product of the probability of whether or not the customer is offered a free installment option or not on any specific purchase occasion times the probability of their choice of installment term (where the choice probability is conditional on whether they are offered a free installment option or not). This would result in a relatively easy estimation exercise, where we could use a flexible parameterization for the option value function and estimate the model no differently than most static discrete choice models are estimated.

In particular, we would then be able to directly observe violations of the weaker version of the dominance assumption, namely we could observe situations where a customer was offered a free installment opportunity of duration $\delta > 2$ and nevertheless, the customer chose a free installment of a shorter duration $d < \delta$. Even though we cannot directly observe such violations of the dominance assumption in our data set, we are able to estimate the probability that they occur, and thereby test the hypothesis that the weaker form of the dominance assumption holds empirically.

However to do this, we need to recognize the difficulties imposed by the fact that our observations of free installment opportunities are *censored* in a way that is very similar to *choice based sampling:* that is, we only observe whether a consumer is offered a free installment option for those purchases where the consumer actually chose the free installment option. In such a situation, how is it possible to infer the probability that customers are offered free installment options? More importantly, how can we estimate the probability that customers do not choose the free installment option when it is offered to them? We show that we can solve the problem by forming a likelihood function that accounts for the censoring. The likelihood function takes the form of a *mixture model* where the probability of being offered a free installment option is a key part of the *mixing probabilities* (there are additional component corresponding to a probability distribution over the duration δ offered to customers who are offered free installment options).

Though there are well known econometric difficulties involved in identifying mixture models, and the degree of censoring in our application is very high (we only observe free installments being chosen in 2.7% of the 167,946 customer-purchase observations used in our econometric analysis), we show that under reasonable but *parametric* assumptions about the forms of the probability function governing free installment options and for flexibly parameterized functional forms for customers' option value functions $ov(a,x,d)$, we are able to separately identify the probability of being offered a free installment, $\Pi(z)$ which depends on a set of variables z including time dummies and merchant class code dummies and consumers' conditional choice probabilities for installments $P(d|a,r,d,x)$.

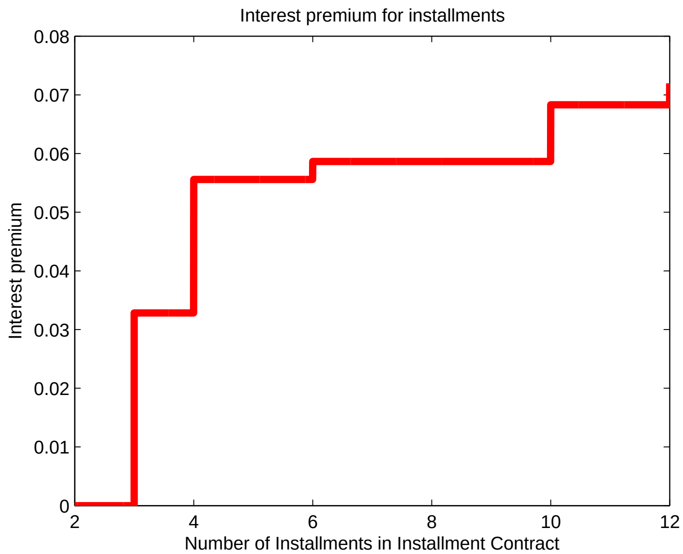
We find that our model fits the data extremely well, but implies a highly inelastic demand for credit. In particular, we find a relatively limited degree of consumer responsiveness to free installment options: the probability of turning down these options is relatively high even though we estimate that for our sample customers are offered free installments approximately 27% of the time, Thus, these customers are taking free installments in only about 15% of the times that they are offered them. We refer to this low take-up rate of what would appear to be a "costless" option for an interest-free loan as the *free installment puzzle.*

Our data are not sufficiently detailed to enable us to delve a great deal further and uncover a more detailed explanation for the reasons *why* customers appear so unwilling to take up free installments and their demand for credit is so inelastic. Our model attributes the reasons for this low takeup rate to a combination of a relatively low option value of credit relative to the cost of credit and to relatively high fixed transactions costs associated in undertaking each installment purchase transaction. However these "transactions costs" could also be interpreted as capturing *stigma* associated with installment transactions, and the low option value may be associated with a fear (whether rational and well-founded or not) that installment credit balances could undermine one's credit rating, or that there are some unspecified hidden future fees or "gotcha's" associated with installment loans beyond the interest rate (e.g. an unfounded belief that there are pre-payment penalties, or a concern that an installment balance could lead to a higher risk of missed future payments and thus late fees). Unfortunately, we are unable to delve further to determine which of these various more subtle psychological explanations is the dominant explanation of the free installment puzzle.

Customers who were not offered interest-free installment purchase options, or who desire a greater number of installment payments than they were offered under an interest-free installment opportunity can borrow (subject to borrowing limits that we do not directly observe in our data) according to a nonlinear, increasing customer-specific interest rate schedules. These schedules are determined according to a rather complex function of a) the consumer's credit score and payment history (including the number of recent late payments), b) the number of installment payments, and c) the current economic environment, including the level of overall interest rates and dummy variables capturing current economic conditions. Though the credit card company does not publish and did not provide us with the formula it uses to set interest rates on installment loans, we were able to uncover it econometrically.

As we described in section 2, we were able to calculate the internal rate of return for each installment loan contract in our data. For the subset of installment contracts where a positive internal rate of return was calculated, we regressed this internal rate of return on the customer specific variables, as well as time and merchant dummies in order to uncover the formula the company uses to set interest rates. Our regression resulted in an extremely good fit, with an R^2 value of 0.99, indicating that we were successful in econometrically uncovering the interest formula the company uses to set interest rates to its customers. We found that the most important factors determining the customer-specific interest rates are factors a) and b) above. In particular, we found that consumer characteristics a) determine the "base interest rate" for

Figure 23: Interest Premium for Installment Purchases as a function of the Installment Term



an installment loan with $d = 2$ payments, but there is a step-wise increasing schedule that is *common to all consumers* that determines successive increases in the interest rate offered for longer installment terms $d > 2$. Figure 23 graphs the interest "premiums" customers must pay for successively longer installment terms d.

Let $\overline{r}(d,x)$ denote the *installment interest rate schedule* offered to a consumer with characteristics x who desires to finance an installment purchase with d installments. By our discussion above, this schedule has the form

$$\overline{r}(d,x) = \rho_0(x) + \rho_1(d), \tag{7}$$

where the characteristics of the particular consumer x only enter via the "intercept" term $\rho_0(x)$, and $\rho_1(d)$ represents the *interest premiums* for installments longer than $d = 2$ months. Thus $\rho_1(d) = 0$ for $d \leq 2$ and $\rho_1(d) > 0$ is given by the function graphed in figure 23 for $d \geq 2$. Note that our regression analysis of actual interest rates charged to customers confirms that the ρ_1 function is, to a first approximation, independent of x and thus common to all of the company's customers.

Consider a consumer with characteristics x who is interested in purchasing a given item that costs an amount a. We take as a given that the consumer is going to make the purchase and focus on modeling the customer's choice of installment term, i.e. whether to pay the balance a in full at the next statement ($d = 1$), or request an installment purchase option with $d > 2$ installments at an interest rate of $r = \overline{r}(d,x)$. Later, we will consider separately the question of how interest rate schedule affect the size of the transaction by estimating the conditional distribution $f(a|x,r,c)$ in equation (1) in section 4.5.

Let $v(d,x,a,r)$ represent the net gain in utility the consumer obtains from choosing installment option d (where again, we have normalized the net gain for paying in full, $d=1$ to $v(1,x,a,r)=0$. Since we do not expect to be able to perfectly predict every consumer's choice of installment term d, we introduce to commonly used device of Type I extreme value unobservable components of utility $\varepsilon(d)$ (unobservable to the econometrician, but not to the customers) that also affect the choice of installment term. We assume that $\varepsilon(d)$ and $\varepsilon(d')$ are independently distributed if $d \neq d'$ and that $E\{\varepsilon(d)\}=0$ for $d \in D$ but with unknown common scale factor $\sigma > 0$ that is an additional parameter to be estimated.

The consumer chooses installment term $d \in D = \{1,2,\ldots,12\}$ if and only if

$$v(d,x,a,\overline{r}(d,x)) + \varepsilon(d) \geq \max_{d' \in D} \left[v(d',x,a,\overline{r}(d',a)) + \varepsilon(d') \right]. \tag{8}$$

The extreme value assumption implies that the conditional probability of observing the consumer choose installment term d is (after integrating out the unobserved components of utility $\{\varepsilon(d')|d' \in D\}$ is given by the standard multinomial logit model

$$P_+(d|a,x) = \frac{\exp\{v(d,x,a,\overline{r}(d,x))/\sigma\}}{\sum_{d' \in D} \exp\{v(d',x,a,\overline{r}(d',x))/\sigma\}}, \tag{9}$$

where the $+$ subscript denotes a choice situation where the consumer can only choose from installment that have positive interest rates, $\overline{r}(d,x) > 0$ for $d \in \{2,\ldots,12\}$. The choice set D in this case is just the set $D = \{1,2,\ldots,12\}$ where choice $d=1$ denotes the decision to pay the amount of the purchase a in full at the next statement date, and choices $d = 2,3,\ldots,12$ denote the decision to spread out the payment over d installments over the next d statement dates, though at the cost of a positive interest rate on the outstanding installment balance.

The consumer's choice problem is slightly more complicated when the consumer is offered an interest-free installment option. Suppose this consumer is offered an interest-free installment option with a maximum duration of δ_0 payments (months) where $\delta_0 \leq 12$. The consumer can either to choose to pay in full, $d = 1$, or purchase the item via the interest-free installment option but over any number of installments $d \in \{2,\ldots,\delta_0\}$, or to pay over even longer installment durations $d \in \{\delta_0+1,\ldots,12\}$, but at the cost of paying a positive interest rate on these installment balances. The consumer will choose a free installment option $d_0 \in \{2,\dots,\delta_0\}$ that satisfies

$$v(d_0,x,a,0)+\varepsilon(d_0) = \max\left[\max_{d\in\{1,\dots,\delta_0\}} v(d,x,a,0)+\varepsilon(d), \max_{d'\in\{\delta_0+1,\dots,12\}} \left[v(d',x,a,\overline{r}(d',a))+\varepsilon(d')\right]\right], \tag{10}$$

The consumer will choose a positive interest rate installment option $d_+ \in \{\delta_0+1,\dots,12\}$ that satisfies

$$v(d_+,x,a,r(d_+,a))+\varepsilon(d_+) = \max\left[\max_{d\in\{1,\dots,\delta_0\}} v(d,x,a,0)+\varepsilon(d), \max_{d'\in\{\delta_0+1,\dots,12\}} \left[v(d',x,a,\overline{r}(d',a))+\varepsilon(d')\right]\right], \tag{11}$$

with the understanding that the set of positive interest rate choices $\{\delta_0+1,\dots,12\}$ is empty if $\delta_0 = 12$. The implied choice probability is denoted by $P_0(d|x,a,\delta_0)$ and is given by

$$P_0(d|x,a,\delta_0) = \frac{\exp\{v(d,x,a,\overline{r}(d,x))/\sigma\}}{\sum_{d_0=1}^{\delta_0}\exp\{v(d_0,x,a,0)/\sigma\}+\sum_{d_+=\delta_0+1}^{12}\exp\{v(d_+,x,a,\overline{r}(d_+,x))/\sigma\}}, \tag{12}$$

if $d \in \{\delta_0+1,\dots,12\}$, i.e. the consumer chooses an installment term longer than the maximum free installment duration offered, δ, or

$$P_0(d|x,a,\delta_0) = \frac{\exp\{v(d,x,a,0)/\sigma\}}{\sum_{d_0=1}^{\delta_0}\exp\{v(d_0,x,a,0)/\sigma\}+\sum_{d_+=\delta_0+1}^{12}\exp\{v(d_+,x,a,\overline{r}(d_+,x))/\sigma\}}, \tag{13}$$

if $d \in \{1,\dots,\delta_0\}$, i.e. the consumer chooses to pay the amount purchased a in full at the next statement date, or chooses one of the free installment options to pay the amount a in 2 to δ_0 installments.

4.2 Likelihood Function

The parameters to be estimated are $\theta = (\sigma,\phi,\alpha,\beta)$ where ϕ are parameters of consumers' utility/value functions $v(d,a,x,r,\phi)$. For notational simplicity, we will include the extreme value scale parameter σ as part of the ϕ vector, so the implied choice probabilities when a consumer is offered a free installment offer of duration δ_0, $P_0(d|a,x,\delta_0,\phi)$, and the choice probability when the consumer is not offered a free installment offer, $P_+(d|a,x,\phi)$, are both functions of an unknown vector of parameters ϕ to be estimated. The parameter subvector α represents parameters characterizing the probability $\Pi(z|\alpha)$ that a customer is offered a free installment offer (where z are variables characterizing the date and merchant category), and β are parameters characterizing the distribution of offered durations of free installment offers $f(\delta_0|z,\beta)$. We use the method of maximum likelihood to estimate these parameters. Below, we describe the likelihood function that accounts for the fact that in certain situations we do not observe whether or not a customer is offered a free installment opportunity.

Consider the likelihood function for a specific customer who makes purchases at a set of times $T = \{t_1, \ldots, t_N\}$. Of these times, there is a subset $T_I \subset T$ where the customer purchased under installment, i.e. where $d > 1$. The complement T/T_I consist of times where the customer purchased without installment, i.e. where $d = 1$. We face a censoring problem that in many cases where $d = 1$, we do not know if the consumer was eligible for an interest-free installment purchase option or not. Even when $d > 1$, we only know if the consumer was offered an interest-free installment purchase option when the customer actually chose that alternative. However it is possible that in some cases customers may have been offered an interest-free installment purchase option with term δ_0 but decided to choose a longer term option at a positive interest rate. Our likelihood must be adjusted to account for these possibilities and to "integrate out" the various possible interest-free installment options that the consumer could have been offered but which we did not observe.

As noted above, $\Pi(z_{it}|\alpha)$ is the probability that a customer i who makes a credit card purchases at date t is offered an interest-free installment opportunity. The vector z_{it} does not contain any customer-specific variables x, but does include dummies indicating the date of the purchase and the type of merchant the customer is purchasing the item from, since as we noted above the main determinants of the interest-free installment option are a) the time of year, and b) the type of merchant (since different merchants can negotiate interest-free installment deals with the credit card company as a way of increasing their sales). Conditional on being offered an interest-free installment purchase option, let $f(\delta_0|z,\theta)$ be the conditional distribution of the installment term that is associated with the interest-free installment option. Note that $f(1|z,\theta) = 0$: by definition an installment payment plan must have 2 or more future payment dates. Equivalently, by default every consumer has the option to pay in a single installment, and they get what amounts to an interest free loan covering the duration between the date of purchase until the next billing date.

Let T_0 be the subset of purchase dates T where the customer did choose the installment option and we observe that this was an interest-free installment option (we can determine this by observing that the consumer never made interest payments on the installments as described above). For this subset, the component of the likelihood is

$$L_0(\theta) = \prod_{t \in T_0} P(d_t|x_t, z_t, a_t, \theta) \tag{14}$$

where

$$P(d|x,z,a,\theta) = \sum_{\{\delta_0|d\leq\delta_0\}} P_0(d|x,a,\delta_0,\phi)\, f(\delta_0|z,\beta)\Pi(z|\alpha), \tag{15}$$

where for each transaction in the set of times T_0, d_t is less than or equal to the free installment (maximum) term $\delta_{0,t}$ offered to the customer under the interest-free installment option and of course $d_t > 1$ (otherwise the consumer would have chosen to pay the amount a_t in full at the next statement date). When the (weak) dominance assumption holds, we have $P_0(d_t|x_t,a_t,\delta_{0,t},\phi) = 0$ if $d_t \in \{2,\ldots,\delta_{0,t}-1\}$, i.e. the customer always chooses the maximal loan duration permitted under the free installment offer. In that case we have $d = \delta_0$ and

$$P(d|x,z,a,\theta) = P_0(d|x,a,d,\phi)\, f(d|z,\beta)\Pi(z|\alpha). \tag{16}$$

Now consider the likelihood for the cases, $t \in T/T_0$, where we do not know for sure if the customer was offered the interest-free installment option or not. There are two possibilities here: a) the consumer chose not to purchase under installment, b) the consumer chose to purchase under installment but paid a positive interest rate, rejecting the free installment offer. Consider first the probability that $d = 1$, i.e. the consumer chose to pay the purchased amount a in full at the next statement date. Let $P(1|x,z,a,\theta)$ denote the probability of this event, which is given by

$$P(1|x,z,a,\theta) = \Pi(z|\alpha)\left[\sum_{\delta_0\in\{2,\ldots,12\}} P_0(1|x,a,\delta_0,\phi)\, f(\delta_0|z,\beta)\right] + [1-\Pi(z|\alpha)]P_+(1|x,a,\phi). \tag{17}$$

The other possibility is that the customer chose to pay under installment for a duration of d months, for $d \in \{2,\ldots,12\}$ but at a positive rate of interest. In the case where $d = 2$, i.e. where the consumer pays a positive interest rate to pay the purchased amount a over two installments, we deduce that the customer could *not* have been offered a free installment opportunity of 2 or more months due to the company's procedures which essentially force the customer into the free installment offer any time then chosen duration is less than or equal to the maximum duration of the free installment opportunity that it offers to the customer. This implies that $P(2|x,z,a)$ is given by

$$P(2|x,z,a,\theta) = [1-\Pi(z|\alpha)]P_+(2|x,a,\phi). \tag{18}$$

The other cases $d \in \{3,\ldots,12\}$ are where the customer chose a positive interest rate installment option but we cannot be sure whether the customer was offered a free installment or not. In this case we have

$$P(d|x,z,a,\theta) = \Pi(z|\alpha)\left[\sum_{\delta_0<d} P_0(d|x,a,\delta_0,\phi)\, f(\delta_0|z,\beta)\right] + [1-\Pi(z|\alpha)]P_+(d|x,a,\phi). \tag{19}$$

The summation term in the formula for $P(d|x,z,a)$ above reflects the company's billing constraint: the customer is not allowed to choose a positive interest installment option d if the customer had been offered a free installment option of duration δ_0 greater than or equal to d. Let $L_1(\theta)$ denote the component of the likelihood corresponding to purchases that the consumer makes in the subset T/T_0, i.e. purchases either that were not done under installment, or which were done under installment but at a positive interest rate. This is given by

$$L_1(\theta) = \prod_{t \in T/T_0} P(d_t|x_t, z_t, a_t, \theta). \tag{20}$$

where $d_t = 1$ if the customer chose to purchase an item at time t without installment, and $d_t > 1$ if the customer chose to purchase via installment, but with a positive interest rate.

The full likelihood for a single consumer i is therefore $L_i(\theta) = L_{i,0}(\theta)L_{i,1}(\theta)$ where $L_{i,0}(\theta)$ is the component of the likelihood for the transactions that the consumer did under free installment offers (or $L_{i,0}(\theta) = 1$ if the consumer had no free installment transactions), and $L_{i,1}(\theta)$ is the component for the remaining transactions, which were either choices to pay in full at the next statement, $d_{i,t} = 1$, or to pay a positive interest rate for a non-free installment loan with duration $d_{i,t} > 1$. The full likelihood for all consumers is then

$$L(\theta) = \prod_{i=1}^{N} L_{i,0}(\theta)L_{i,1}(\theta). \tag{21}$$

4.3 Model Specification

We maximize the log-likelihood with respect to θ for various "flexible functional forms" for $v(d,x,a,r)$ that are designed to capture the net "option value" to the customer of purchasing an item under installment. We assume that $v(d,x,a,r)$ has the additively separable representation given in equation (6) above. Thus, we can view consumers as making "cost-benefit" calculations where they compare the benefit or option value $ov(a,x,d)$ of paying a purchase amount over $d > 1$ installments with the interest costs $c(a,r,d)$. For free installments, we have $c(a,r,d) = 0$, but this does not necessarily imply that customers will necessarily always take every free installment option. One reason is due to the randomly distributed *IID* extreme value shocks $\varepsilon(d)$ representing unobserved idiosyncratic factors that affect a consumer's choice of the installment term. In some cases these shocks will be sufficiently negative to cause a consumer not to take a free installment offer even if $ov(a,x,d)$ is positive (and thus higher than the utility of paying the purchase in full at the next statement date, which is normalized to 0). Another reason is that we specify the option value function as follows

$$ov(a,x,d) = a\rho(x,d) - \lambda(x,d) \tag{22}$$

where we can think of $\rho(x,d)$ as the percentage rate a customer with characteristics x is willing to pay for a loan of duration d months and $\lambda(x,d)$ represents the fixed transaction costs of deciding and undertaking an installment transaction at the checkout counter. Note that this component is assumed not to be a function of the amount purchased a whereas the other component of the option value, $a\rho(x,d)$ is a linear function of the amount purchased. This implies that *consumers will not want to pay for sufficiently small credit card purchases on installment since the benefit of doing this,* $a\rho(x,d)$*, is lower than the transactions cost* $\lambda(x,d)$. We can also think of λ as capturing potential "stigma costs" associated with purchasing on installment, as well as "mental accounting costs" such as any apprehension customers might have that adding to their installment balance increases their risk of making a late payment on their installment account in the future, or that undertaking another installment transaction will have adverse effects on their credit score, and so forth.

Notice that we assume the option value of having the benefit of extended payment does not depend on the interest rate the credit card company charges the customer, and the customer-specific interest rate schedule $\overline{r}(d,x)$ only enters via the cost function $c(a,r,d)$. This is an important identifying assumption. Furthermore we assume that the financial cost that a customer perceives due to purchasing an item under installment equals the excess of the total payments that the customer makes over the term of the agreement less the current cost a of the item. That is, we assume c equals the difference between the total payments the customer makes under the installment agreement *cumulated with interest to the time the installment agreement ends* less the amount the customer purchased, a, discounted back to the date t when the customer purchased the item. This value can be shown to be

$$c(a,r,d) = a\left(1 - \exp\{-rt_d/365\}\right), \tag{23}$$

where t_d is the elapsed time (in days) between the next statement date after the item was purchased and the statement date when the final installment payment is due. The interest rate r is the internal rate of return on the installment loan, and is given by $r = \overline{r}(d,x)$. Recall that this is the positive interest rate that company offers to the customer for an installment purchase with term d. Notice that if $d = 1$ and the consumer chooses not to do an installment then $c(a,r,1) = 0$. Notice also that for any interest-free installment opportunity, $r = 0$ and so $c(a,r,d) = 0$ as well. To a first approximation (via a Taylor series approximation of the exponential function) we have $c(a,r,d) = rat_d/365$, so the cost of the installment loan equals the product of the duration of the loan, the amount of the loan, the interest rate offered to the consumer, times the fraction of the year the loan is outstanding.

Notice that the $c(a,r,d)$ function has no unknown parameters to be estimated. The parameters to be estimated are the parameters ϕ entering the option value function, $ov(a,x,d,\phi)$, the scale parameter σ of the Type I extreme value distributions for the unobserved components of the $v(a,x,r,d,\phi)$ functions, and α, the parameters entering the probability of being offered a free installment, $\Pi(z)$ and the probability distribution over the maximum term of the free installment offers that are offered to consumers, $f(d|z,\beta)$. Recall that $\theta = (\sigma,\phi,\alpha,\beta)$ is the full set of parameters to be estimated. Table 3 presents the maximum likelihood estimates of (σ,ϕ). We discuss the maximum likelihood estimates of the 26 α parameters later. Clearly, the parameters of interest are (σ,ϕ). We are not interested in the α parameters *per se,* though we do want to know if our estimate of the conditional probability $\Pi(z,\alpha)$ of receiving a free installment offer is reasonable.

To understand the parameter estimates, note that we have specified $ov(a,x,d) = a\rho(x,d)$ where

$$\rho(x,d) = \frac{1}{1+\exp\{h(x,d,\phi)\}} \tag{24}$$

where

$$\begin{aligned} h(x,d,\phi) \quad &= \quad \phi_0 I\{d \geq 2\} - \sum_{j=3}^{12} \exp\{\phi_{j-2}\} I\{d \geq j\} + \phi_{11}\text{ib} + \phi_{12}\text{installshare} \\ &\quad +\phi_{13}\text{creditscore} + \phi_{14}\text{nlate} + \phi_{15} I\{r=0\}. \end{aligned} \tag{25}$$

The fixed transaction cost of choosing an installment term at the checkout counter, $\lambda(x,d)$, is specified as

$$\lambda(x,d) = \exp\left\{\phi_{16} I\{r=0\} + \phi_{17}\text{installshare} + \sum_{j=2}^{10} \phi_{16+j} I\{d=j\} + \phi_{27} I\{d>10\}\right\}. \tag{26}$$

The variable *creditscore* is the interpolated credit score for the customer at the date of the transactions (the company only periodically updates its credit scores so we only observed them at monthly intervals), and *nlate* is the number of late payments that the customer had on his/her record at the time the transaction was undertaken, and *ib* is the customer's installment balance at the time of the transaction. Note that due to the large variability in spending on credit cards by different customers, we normalized both *a* and *ib* as *ratios of each customer's average statement amount.*

The most important variable of the *x* variables turned out to be *installshare,* the share of creditcard spending that the customer does under installment. We included *installshare* because, as we showed in section 2, it serves as an important observable indicator of unobserved preference heterogeneity, as well as an observed indicator about which consumers are most likely to be liquidity constrained. We found that neither *creditscore* nor *nlate* are as powerful as the *installshare* variable in enabling the model to fit the data and capture the large degree of customer-specific heterogeneity that we observe in our sample.

An alternative strategy would be to replace *installshare* by a random parameter τ representing *unobserved heterogeneity* with the interpretation that lower values of τ indicate customers who are more desperate for liquidity and thus have a higher subjective willingness to pay for loans of various durations, $\rho(x,d,\tau,\phi)$. However, we have had considerable difficulty so far in estimating specifications with unobserved heterogeneity due to the fact that we have an unbalanced panel where for some consumers we observe many hundreds of transactions. Conditioning on τ, the likelihood for these hundreds of conditionally independent choices of installment duration is typically a *very very small number.* Unobserved heterogeneity specifications require us to take averages (i.e. integrate over the distribution of τ) of these very small numbers and we often found that when we tried to take the logarithm of the resulting *mixture probability* it was sufficiently small to be below the "machine epsilon" i.e. the lowest positive number a computer is capable of representing, even on 64-bit machines.

In view of these problems, we found *installshare* to be extremely convenient as an "observed indicator" of the underlying unobserved heterogeneity τ. We conjecture that if we can somehow resolve the problem of "underflow" in computing the mixing probabilities, the estimation results (particularly the overall fit of the model) of a specification with a sufficiently rich specification of unobserved heterogeneity but omitting *installshare* will be quite similar to the results presented below with τ omitted and *installshare* included.

4.4 Identification

It is not immediately obvious that the model we specified in sections 4.2 and 4.3 above is identified. Even without accounting for the mixture model specification that results from accounting for unobserved heterogeneity as described in section 4.3, the likelihood function we derived in section 4.2 can already be regarded as a type of *mixture model* since the conditional probabilities $P(d|x,z,a,\theta)$ entering the likelihood function are themselves mixtures of the underlying choice probabilities $P_0(d|x,a,\delta,\phi)$ and $P_+(d|x,a,\phi)$ that constitute the probabilities of choosing different installment terms with and without the presence of a free installment offer with maximum duration δ, respectively. As is well known, it is very difficult to identify econometric models that are formulated as mixtures of probabilities, since a wide variety of probability distributions can be well-approximated by convex combinations of a given a set of probabilities (also known as "components"), and there are generally many different ways to do this. For example, Henry et al. [2011] note that "Without further assumptions there is of course no way to identify the mixture weights and components" (p. 2).

Identification can be especially problematic when we relax the weak dominance assumption, since then both of the conditional probabilities P_+ and P_0 have the same support $\{1,\dots,12\}$, and the conditional probabilities entering the likelihood are mixtures of these two conditional probabilities. If we view the identification problem from the lens of "multicollinearity", another way to state the concern about identification is that it is far from obvious that probabilities P_0 and P_+ are sufficiently different from each other to rule out the possibility that are many different ways to represent the "reduced-form" probabilities $P(d|x,z,a,\theta)$ that enter the likelihood in terms of various convex combinations of the "structural" probabilities P_+ and P_0.

The identification of our model is key to the plausibility of the conclusions we draw about individual behavior from this exercise. To see why, consider the following two explanations for the relatively small fractions (2.7%) in our sample that are done as free installments: a) consumers will take virtually any free installment that it is offered to them (so the strong dominance assumption holds and $P_0(\delta|x,a,\delta,\phi)=1$) and the average probability of being offered a free installment is very low (i.e. about 2.7%), versus explanation b) the average probability of being offered a free installment is very high, but consumers are averse to choosing free installments, so that even though the probability of being offered a free installment is high, the probability that it is chosen is sufficiently low that the average probability that free installments are actually offered *and* are chosen is very small, i.e. approximately 2.7% on average. It is not obvious how the method of maximum likelihood can distinguish between these two competing explanations for the low share of free installments in our sample.

Despite these concerns, we find that our model *is* identified and surprisingly, the method of maximum likelihood is able to distinguish between the two explanations a) and b) for the low take up rate of free installments, with the likelihood for hypothesis b) being sufficiently greater than the likelihood of hypothesis a) that we are easily able to reject a) in favor of b). Note that our model is fully *parametric* and the standard argument for identification of parametric involves showing that the expectation of the log-likelihood function, $E\{\log(P(\tilde{d}|\tilde{x},\tilde{z},\tilde{a},\theta))\}$ is uniquely maximized at a value θ^* in the parameter space.

As is well known, in the case of the multinomial logit model, the expectation of the log-likelihood is *concave* in the underlying parameters, and identification amounts to verifying additional conditions that imply that this function is also *strictly concave.* However the concavity property generally no longer holds when the expected log-likelihood function involves mixtures of multinomial logit models. When a parametric model is unidentified, there are typically two ways in which the identification condition fails: either 1) the expected log-likelihood function is "flat" in a neighborhood of the global maximum (so there is a continuum of values of θ that maximize the likelihood), or 2) each local maximum of the expected log-likelihood is "regular" in the sense that the hessian matrix at each local maximum is negative definite (implying that there are a finite number of isolated local maxima, each one is unique within a sufficiently small neighborhood of each local maximum point) but there are two or more distinct local maxima that happen to have the same exact value of the expected log-likelihood, so the set of such distinct global optima are observationally equivalent and the model is unable to distinguish them.

Given the large number of observations in our sample, $N = 167,946$, the empirical log-likelihood $\log(L(\theta))/N$ (where $L(\theta)$ is the likelihood function defined in equation (21) above) provides a very good approximation to its expectation $E\{\log(P(\tilde{d}|\tilde{x},\tilde{z},\tilde{a},\theta))\}$ by the uniform law of large numbers. Therefore it is sufficient to show that the sample log-likelihood function has a unique maximizer since for the very large sample size we have in this case, the probability is very high that the sample log-likelihood is uniformly close to its expectation. Therefore, since the hessian of the likelihood is a continuous function of the parameters θ, the continuous mapping theorem implies that if the sample log-likelihood has a unique maximizer (or equivalently each local maxima that we find are "regular" — the type 2 case discussed above), then we can rule out the most obvious type of non-identification, i.e. namely that the expected log-likelihood is locally flat in a neighborhood of the global maximum. We have indeed verified this numerically: at each local maximum we found in the course of a thorough search of the likelihood over the parameter space, we found that the hessian of the sample log-likelihood function was negative definite.

Further, though we did encounter multiple local maxima of the likelihood function in the course of running our estimation algorithm, we we unable to find distinct local maximizers that resulted in the *identical* values of the sample log-likelihood function. Instead we found a single "global optimum" $\hat{\theta}$ that resulted in a significantly higher sample log-likelihood than for any of the local optima we encountered in our thorough search for a global optimum of the likelihood. Although we are not aware of any general argument that we can rely to provide a mathematical proof that there are no other values of θ besides the value we found $\hat{\theta}$ that result in the same or a higher value of the sample log-likelihood function, we feel that our numerical experience in maximizing the likelihood does at least provide strong evidence suggesting that the parameters of our model are in fact identified.

The intuition for how the data are able to distinguish between the two hypotheses a) and b) for the low take up of free installments discussed above is as follows. If hypothesis a) were the correct one, i.e. that the strong dominance assumption holds (or nearly holds), then consumers would take nearly every free installment opportunity that is offered to them and the low incidence of free installments could only be a result of free installments being rarely offered to consumers. However in this case, the model would assign a high option value to borrowing under installment — at least sufficiently high that consumers' option values exceed any fixed costs involved in undertaking the free installment transaction. However the high option value would then imply that customers who have sufficiently low positive of positive interest installments should also be frequent users of installment credit, something we do not observe in our data. This provides an intuitive argument for how the data are able to reject the strong dominance assumption and instead provide strong evidence in favor of explanation b) as the model most consistent with the data we observe.

Identification of the parameters β is assisted by an assumption we made that the merchant/date variables in z do not affect the distribution of free installment durations, so we write this distribution as $f(d|\beta)$ rather than as $f(d|z,\beta)$. This assumption was motivated out of concerns that $f(d|\beta)$ would be difficult to identify that the probability of receiving a free installment offer itself, $\Pi(z,\alpha)$, since when we relax the strong dominance assumption, if we observe a customer taking a free installment offer of duration d the customer could have been offered a free installment with a maximum duration δ for any $\delta \in \{d,\ldots,12\}$. This gives considerable freedom to how the model might "explain" the particular set of installment durations that consumers actually choose. For example, one possibility is to set $f(12|z,\beta) = 1$, so that the maximum duration of every free installment offer is 12, and the pronounced peak we observe in free installments at a duration of $d = 3$ is purely a result of consumers pre-commiting and choosing their most popular loan duration $d = 3$ rather than choosing the full $\delta = 12$ month loan duration. Although this explanation might seem a bit implausible on its face, recall figure 3, which showed that $d = 3$ is the most likely term of installment loan for individuals who choose to do installments at a positive interest rate.

Though we have independent evidence that in fact most free installment loans that are offered to consumers have a maximum of $\delta = 3$ installments, how can the likelihood distinguish between the case where all free installments offered have a maximum of $\delta = 12$ installments versus the case where all free installments have a maximum of $\delta = 3$ installments? One easy way that the latter hypothesis can be rejected is by virtue of the fact that we do observe a small number of free installments that did involve 12 payments. This enables us to conclude that not all free installment offers could have a maximum of $\delta = 3$ installments. However, beyond, this, the precise identification of the probabilities $f(d|z,\beta)$ seems more tenuous, since due to the censoring, we never directly observe someone being offered a free installment with a maximum of δ installments and choosing to take the installment for $d < \delta$ installments.

We do note that we made several implicit *exclusion restrictions* that assist in the identification of the parameters of the model. First, we assume that the z variables that affect the probability of being offered a free installment opportunity do not enter the choice probabilities P_+ and P_0. This is because z contains dummy variables for merchant codes and calendar time intervals that are relevant for predicting whether a free installment is offered but do not seem directly relevant for predicting a consumer's choice of installment term. Conversely, the customer specific variables x do enter these choice probabilities but can be plausibly excluded from the probabilities that a customer would be offered a free installment opportunity. Finally, we also assume that the probabilities of being offered free installments of various maximum durations are independent of z, so only 10 parameters are necessary to estimate these 11 probabilities. Following our pragmatic approach to identification, we verified numerically that various convex combinations of the choice probabilities P_0 (where the duration probabilities $f(d|\beta)$ are the mixture weights) do not result in the same reduced-form probability $P(d|x,z,a,\theta)$. Otherwise the likelihood function would be flat in a neighborhood of any optimum, and this in turn would imply that the log-likelihood function has a singular hessian matrix at any such point. However we found in fact that the hessian is strictly negative definite at the maximum. Further evidence is provided by the fact that if we fix the β parameters at arbitrary values and maximize over the remaining parameters (ϕ,α), the value of the likelihood falls significantly below the value we attain when we free up β and optimize over (ϕ,α,β) simultaneously.

In summary, the identification of our model results from a combination of 1) *exclusion restrictions* and 2) *parametric functional form assumptions.* We have not investigated conditions under which the "structural objects" in our model $\{P_+,P_0,\Pi,f\}$ are *non-parametrically identified* however recent work by Henry et al. [2011] and others may represent promising avenues for further investigation. For this study, we feel that are exclusion restrictions are well-justified and our specification of the option value function ρ and fixed cost functions λ are sufficiently flexible that none of our conclusions are fragile, or depend on arbitrary or hard to justify assumptions. In a fundamental sense, we view the data are telling us that we can separately identify these various probabilities, so the inferences we draw are unlikely be a artifacts of strong, "tricky" modeling assumptions.

4.5 Estimation Results

The estimation results are presented in table 3. Note that in general, most though not all of the parameters are estimated very precisely — something we would expect given the large number of observations in our sample. Due to the large number of α parameters (26) and because they are not of central interest to this paper, we omit them from table (3). However we note that the estimated probabilities of receiving a free installment offer $\Pi(z,\hat{\alpha})$ vary rather significantly over our sample, from a low of 1.41×10^{-4} to a high of 0.527. Over our entire sample, the average estimated probability that a given transaction was subject to a free installment offer is 17%. This estimate appears to be reasonable from our discussions with the credit card company executives. As we see below, it implies that the "take up rate" of free installments is low: although the model predicts substantial consumer-specific heterogeneity in take up rates, on average only 15% of the individuals who are offered free installment opportunities actually take them.

The free installment probabilities vary over the calendar year and across merchants, and the combination of merchant and time dummies enabled us to capture the high degree of variability of free installment options, both over time and across merchants. The variability also justifies our treatment of free installments as "quasi random experiments" since there appears to be no easy way to predict when and where free installments will be offered to consumers.

We now turn to the parameters of interest, the ϕ parameters entering the option value function $\rho(x,d,\phi)$ and the fixed cost function $\lambda(x,d,\phi)$ that are two key "behavioral objects" underlying our discrete choice model. Note that due to the large variability in spending across different consumers, we normalized each customer's credit card spending and installment balances to be ratios of their average statement amounts (the monthly balance due on their credit card bill). Thus, a purchase amount $a=2$ denotes a purchase that is twice as large as the average amount of that customer's average credit card balance on each statement date. An installment balance, denoted as *ib*, equal to 3 would denote an installment balance that is 3 times as large as the average of the customer's credit card balance due.

Consider first the estimation results for the parameters entering the option value function $\rho(x,d,\phi)$. We did not include a constant term in our specification in equation (25) since the sum of the installment duration dummy variables $I\{d\geq j\}$, $j=2,\ldots,12$ adds up to the constant term on the set of relevant

Table 3: Maximum Likelihood Parameter Estimates, Dependent variable: chosen installment term, d

$\rho(x,d,\phi)$ (option value)	Estimate	Standard Error
σ	0.066	3.97×10^{-4}
$\phi_0\ I\{d\geq2\}$	-3.693	0.025
$\exp\{\phi_1\}\ I\{d\geq3\}$	0.227	0.018
$\exp\{\phi_2\}\ I\{d\geq4\}$	0.251	0.179
$\exp\{\phi_3\}\ I\{d\geq5\}$	0.067	0.049
$\exp\{\phi_4\}\ I\{d\geq6\}$	0.136	0.026
$\exp\{\phi_5\}\ I\{d\geq7\}$	2.265×10^{-25}	0.072
$\exp\{\phi_6\}\ I\{d\geq8\}$	4.430×10^{-14}	0.092
$\exp\{\phi_7\}\ I\{d\geq9\}$	0.156	0.079
$\exp\{\phi_8\}\ I\{d\geq10\}$	0.082	0.053
$\exp\{\phi_9\}\ I\{d\geq11\}$	9.070×10^{-15}	0.180
$\exp\{\phi_{10}\}\ I\{d=12\}$	0.281	0.180
ϕ_{11} (ib)	-0.087	0.001
ϕ_{12} (installshare)	-2.202	0.040
ϕ_{13} (creditscore)	-0.207	0.005
ϕ_{14} (nlate)	-0.015	0.002
$\phi_{15}\ (I\{r=0\})$	-2.166	0.061
$\lambda(x,d,\phi)$ (fixed cost)	Estimate	Standard Error
ϕ_{16} (installshare)	-0.941	0.015
$\phi_{17}\ (I\{r=0\})$	-0.246	0.011
$\phi_{18}\ (I\{d=2\})$	-0.740	0.010
$\phi_{19}\ (I\{d=3\})$	-1.006	0.009
$\phi_{20}\ (I\{d=4\})$	-0.297	0.016
$\phi_{21}\ (I\{d=5\})$	-0.487	0.012
$\phi_{22}\ (I\{d=6\})$	-0.208	0.018
$\phi_{23}\ (I\{d=7\})$	-0.106	0.024
$\phi_{24}\ (I\{d=8\})$	-0.106	0.022
$\phi_{25}\ (I\{d=9\})$	-0.462	0.012
$\phi_{26}\ (I\{d=10\})$	-0.215	0.014
$\phi_{27}\ (I\{d>10\})$	-2.166	0.061
$f(d,\beta)$ (maximum installment term)	Estimate	Standard Error
$f(2,\beta)$	0.695×10^{-15}	0.003
$f(3,\beta)$	0.594	0.290
$f(4,\beta)$	1.717×10^{-12}	0.025
$f(5,\beta)$	5.362×10^{-13}	0.022
$f(6,\beta)$	1.356×10^{-14}	0.044
$f(7,\beta)$	3.314×10^{-14}	0.112
$f(8,\beta)$	2.358×10^{-16}	0.150
$f(9,\beta)$	1.565×10^{-11}	0.108
$f(10,\beta)$	0.256	0.425
$f(11,\beta)$	3.252×10^{-16}	0.436
$f(12,\beta)$	0.149	0.024
Log-likelihood, number of observations	$\log(L(\theta))=-46561.3$	$N=167,946$

choices, $d \in \{2,\ldots,12\}$ since we have normalized the option value for the decision $d = 1$ to equal zero. Therefore, we allowed the parameter ϕ_0 to unconstrained and take positive or negative values in order to to play the effective role of the constant term. However we did constrain the coefficients of $I\{d \geq j\}$ for $j = 3,\ldots,12$ to be positive by expressing these as exponential functions of the underlying parameters ϕ_j, $j = 1,\ldots,10$.[4] It is easy to seee that this is equivalent to constraining the option value function $\rho(x,d,\phi)$ to be non-decreasing as a function of d.

Figures 24 and 31 plot the estimated option value function and compares it to the $c(a,r,d)$ function (which, recall, has no unknown parameters in it). However the $c(a,r,d)$ function does depend on the set of interest rates, $r(x,d)$, which do depend on customer characteristics x. We plotted these figures for an illustrative consumer with a creditscore of 2, $ib = 2$, and an installment share of 30%. From figure 24 we see that indeed, the estimated ρ function is non-decreasing in d and it is everywhere above the cost of credit function $c(a,r,d)$, signaling a clear net benefit of purchasing under installment credit. The $\rho(x,d,\phi)$ function has its largest jumps at $d = 3$ and $d = 12$.

Figure 25 plots the net benefits from installment borrowing, $\rho(x,d,\phi) - c(a,r,d)$, as a bar-plot. We see that for this particular customer, the highest net benefits occur at a duration of $d = 4$, where the customer experiences a net benefit to taking an installment, net of the cost of the installment, of about 7% of the transaction amount a. The net benefit of installments is generally the highest for shorter duration installment loans, for $d \in \{2,\ldots,6\}$, and then falls for the longer duration loans $d \in \{7,\ldots,11\}$ but increases again for loans with $d = 12$ installments. This pattern of net benefits is generally consistent with the pattern of installment loan choices, although it does not show any pronounced peak at $d = 3$ that could explain the peak in installments at this duration that we observed in figure 23. We will explain how the model is able to capture this peak when we describe the estimation results for the λ function below.

Other points to note about the estimated parameters of ρ is that the option value *increases* with the size of the customer's existing installment balance (see ϕ_{11} the coefficient of *ib*). The option value is also an increasing function of *creditscore* which means customers with worse (i.e. higher) credit scores are predicted to have higher option values for installment credit. Similarly, another indicator of credit problems, the number of late payments that the customer has on his/her record, *nlate*, also increases the option value and thus the value of installment credit.

[4] In table 3 we report the exponentiated values instead of the parameters themselves, and used the delta method to calculate the implied standard errors.

Figure 24: Estimated option value $\rho(x,d,\phi)$ function relative to $c(a,r,d)$ function

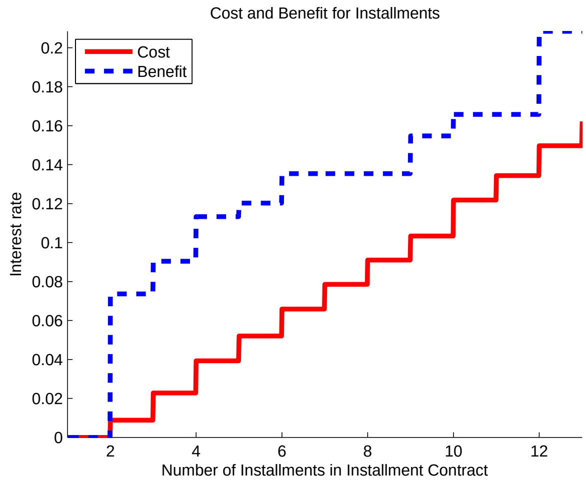


Figure 25: Net benefit of installment Credit as a function of installment duration d

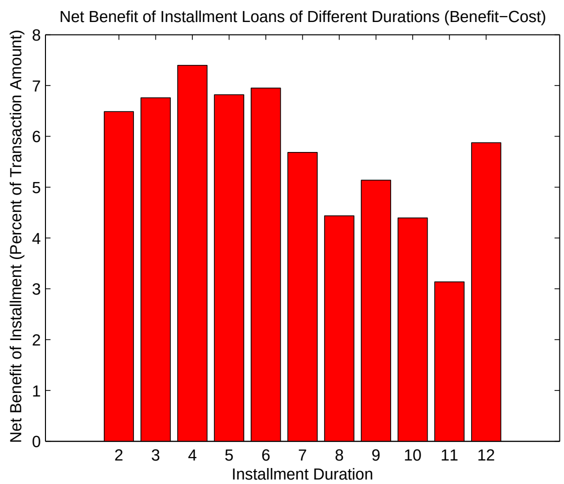
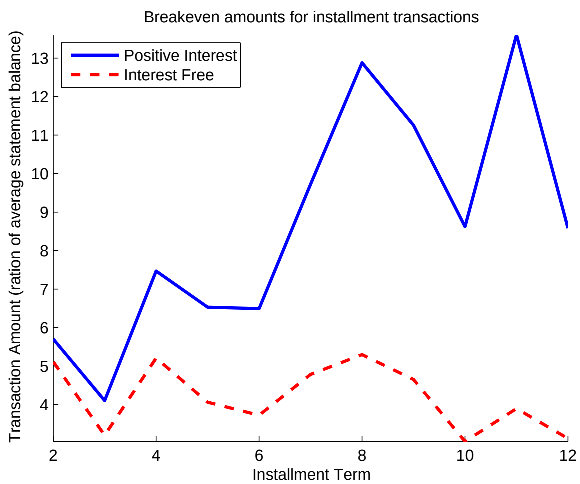


Figure 26: Estimated breakeven amounts $\overline{a}(x,d)$ for installment transactions



The two largest (in absolute value) coefficients after ϕ_0 are ϕ_{12} the coefficient of the *installshare* variable, and ϕ_{15}, the coefficient of a dummy variable indicating that the transaction was done as a free installment. The latter coefficient indicates that customers perceive free installments to have even *higher* option value than installments done at positive interest rates. We are not quite sure of how to interpret this finding, but the data are clearly telling us that it needs to provide an extra boost to the option value in order to explain the take up rate of free installment opportunities. Perhaps one explanation could be that consumers enjoy the value of a loan that much more when they know it is a free loan. This tells us that our specification of cost function $c(a,r,d)$ and our formulation of the installment loan as a simple cost-benefit tradeoff is not sufficient not fully capture how consumers evaluate free installment offers.

Finally, the negative and strongly statistically significant estimated coefficient of the *installshare* variable ϕ_{12} indicates, not surprisingly, that customers with high installment shares have uniformly higher estimated option values, and thus a higher proclivity to take installments, whether free installments or at positive interest rates. As we discussed previously in section 4.3, we used the *installshare* as an observable indicator of unobserved heterogeneity, since we found it infeasible to implement a random effects approach to control for unobserved heterogeneity for the reasons already discussed in section 4.3. We view the *installshare* variable as capturing customers who are "credit constrained" in ways that are not well captured by the *creditscore* and *nlate* variables, though it may also capture customers who are for some other reason "installment addicts" who make frequent use of installment credit. Some of these could be consumers who behave like the textbook *homos economicus* with time-separable utilities and non-hyperbolic geometric discounting of future utilities that result in time-consistent intertemporal preferences and the prediction that these individuals would never pre-commit *ex ante* to choices that reduce their future borrowing options, at least without any obvious compensation for doing so.

We now turn to a discussion of the estimated parameters of the fixed cost function $\lambda(x,d,\phi)$. Generally, the model estimates indicate that consumers perceive high fixed costs to choosing any installment transactions other than the "default" choice $d = 1$. These "costs" may reflect perceived "stigma" associated with taking installment transactions. From anecdotal evidence, the people in the country we are studying regard installment purchases as a sign of "weakness" especially in view of the bad experience that these people had several years prior to the period we studied where there had been a credit bubble and a high frequency of credit card defaults. Thus, the individuals may have been chastised or even scarred by that prior experience and had resolved themselves to try to avoid the use of installment credit whenever possible.

One might ask why this scarring effect and aversion to installments doesn't show up in lower estimated option values. We believe that the fixed costs play an important role in explaining a clear pattern in our data where generally only sufficiently expensive purchases are made under installment. Recall figures 12 and 12 which showed that while the average credit card purchase is $74, the average installment purchase is $364, or nearly 5 times larger than the average credit card purchase. The fixed costs are estimated to be large in order to explain this differential pattern of spending.

Figure 26 illustrates this by plotting the "cut-off" value of spending $\overline{a}(x,d)$ for which the net benefit of borrowing on installment equals the fixed cost of undertaking it, i.e.

$$\overline{a}(x,d) = \frac{\lambda(x,d,\phi)}{\rho(x,d,\phi) - c(a,r(x,d),d)}. \tag{27}$$

This figure was calculated for an individual with a *creditscore=5* (i.e. about average credit) with *installshare=.1* and $ib = 0$ and *nlate=4*. We see that for positive interest loans, the breakeven ratio (i.e. the amount is expressed as a ratio of the average credit card statement balance) is generally over 5 and is as high as 12 or 13 for the less popular installment loan durations, $d = 8$ and $d = 11$.

Notice that ϕ_{17}, the coefficient of $I\{r = 9\}$ is *negative and strongly statistically significant* indicating that consumers perceive free installments to have lower fixed costs, even though at the same time they perceive the option value for free installment loans to be lower as well. Again, we are not quite sure how to interpret this, but one possible interpretation is that since the free installment is a promotion, the merchant may arrange extra assistance by the checkout clerk or provide other cues to try to encourage customers to take the free installment, and this might show up in our model as a lower cost for choosing a free installment loan over a comparable installment loan at a positive interest rate.

The net effect of free installment offers on credit decisions is to lower the $\overline{a}$ threshold since we already showed that the free installment offer increases the option value of the loan, and it also zeros out the cost of the loan which increases the denominator of (27), and it also reduces the fixed costs of taking an installment loan are estimated to be lower if the loan is a free installment offer, and this reduces the numerator of (27). This effect is illustrated in figure 26 for the particular customer that we plotted, and is particularly pronounced for loans of duration $d = 8$ and higher: under a free installment offer the cutoff point is less than 5 and as low as 3 times their average statement amount, whereas the cutoffs are over 10 for positive installment loans. We believe this effect explains the counterintuitive finding of section 2, where we showed that the average free installment loan amount was *lower* than the average positive interest loan amount. This is also what our estimated model predicts as well, and we believe it explains the counterintuitive findings from the matching estimator in table 2 of section 3. Even though it is true free installment offers tend to significantly lower the threshold at which a customer is likely to accept a free installment offer, thereby *reducing* the mean size of a free installment transaction relative to a positive interest installment transaction, we cannot conclude from this that the demand for credit is upward sloping. In fact, we show below that our model predicts that the demand for credit is downward sloping, even though it also predicts this counter-intuitive reduction in average amounts purchased under free installment offers.

The final comment we have about the estimated λ function is that the coefficient ϕ_{16} of the *installshare* variable is a large negative number that is very precisely estimated. Thus, we find that the model captures the systematically higher use of installment credit by individuals with high values of *installshare* by increasing the option value of the loan and by reducing the fixed cost of undertaking the transaction. This is how the model explains our finding in figure 17 of section 2 that the ratio of the typical installment purchase to the typical credit card (non-installment) purchase decreases as *installshare* increases.

Finally, we discuss the estimated probabilities $f(d|\beta)$ representing the probability distribution over the maximum duration of a free installment offer, conditional on one being offered to a given customer. Recall that in section 4.3 we discussed concerns about our ability to identify this probability distribution with much precision. We see that fortunately, the estimation does not imply that all free installment offers involve a maximum of $\delta = 12$ installments, something we know is not the case from our discussions with the credit card company. Instead, the estimation results are very reassuring, since they show that the most commonly offered installment is for a maximum duration of 3 installments, something that we also believe is the case from discussion with executives of the credit card company. However we were surprised to see that the point estimates of our model imply that there is a near zero probability of being offered a free installment for a duration of $\delta = 6$ months.

The difficulty of identifying the $f(d|\beta)$ probabilities is indicated by the large estimated standard errors relative to the point estimates (again, the standard errors for $f(d|\hat{\beta})$, $d \in \{2,\dots,12\}$ were computed from the standard errors for $\hat{\beta}$ using the delta method). The large standard errors reflect the uncertainty our model has in estimating these probabilities even with $N = 167,946$ observations. Given these large standard errors, there does appear to be a fairly wide range of distributions $f(d|\beta)$ that could be consistent with the installment choice data we observe. However these probabilities are not of direct interest to us in this study: instead, we are interest in consumer behavior and the uncertainty in the estimated β coefficients fortunately does not transmit and result in huge uncertainty in the key ϕ parameters entering the ρ and λ

functions. As a result, we are confident that our inferences and key behavioral conclusions are robust to our uncertainty about the probabilities $f(d|\beta)$.

We conclude this subsection with a discussion of our estimation results for the parameters of the distribution of purchases $f(a|x,r,c)$ that enters the expected demand curve for installment credit in formula (1) of section 3. Via initial non-parametric estimation for various consumers, we found that this distribution is well approximated as a log-normal probability density, so we estimated its parameters via regression using $\log(a)$ as the dependent variable. However for the reasons expressed above we were concerned about potential endogeneity in the consumer-specific interest rates. Therefore we conducted a series of regressions, using various types of fixed-effect regressions (e.g. regressing first differences of $\log(a)$, or $\log(a)$ less customer-specific sample means of $\log(a)$, or estimating customer-specific intercepts, etc.) that are possible given the panel nature of our data and the fact that we observe many purchase transactions for each customer in our data set. We found that regardless of how we accounted for fixed effects and whether we did OLS or instrumental variable regressions (where similar to our regressions in section 3 we used the CD rate as an instrumental variable for r) the estimated coefficient of r is extremely sensitive to the inclusion of time dummy variables in our regression. When time dummies are included, the coefficient of the interest rate is estimated to be near zero with a large standard error, allowing us to easily reject the hypothesis that r affects purchase amounts.

However when we omit the time dummies, then the coefficient of r is estimated to be negative and statistically significant in our two stage least squares regressions. However we do not believe this latter result is the correct one. Note that we have relatively few customer-specific variables x, and thus, the regression has no good way to account for macroeconomic shocks that affect credit card spending other than via the interest rate, which typically moves countercyclically. Thus, in in good times interest rates tend to be high and credit card spending tends to be high, whereas in bad times interest rates tend to be low and credit card spending is lower too. This suggests that interest rates should be *positively correlated* with credit card spending, however as we discussed in section 3, we also find that our instruments, such as the CD rate, is negatively correlated with customer-specific interest rates. As a result, the two stage least squares regression predicts a weak negative relationship between the instrumented consumer-specific interest rate and credit card spending.

However in the absence of adequate explanatory variables for income, employment, and other factors that have strong direct effects on household spending decisions, including credit card spending, we believe that time dummies are a next best substitute for capturing macroeconomic shocks that affect all households. Thus, when we include these time dummies, the estimated coefficient on the interest rate in our regressions falls to near zero and has a very large estimated standard error. Our conclusion is that it is plausible that credit card interest rates have negligible direct impact on credit card spending decisions, especially given that the vast majority of transactions in our sample are done without the benefit of any installment credit. In any event, we feel that the data at our disposal is not sufficiently rich in customer-specific covariates that we think are likely to have much stronger effects on credit card spending decisions than interest rates (such as family income, employment, and other unexpected spending shocks such as health shocks and so forth) that we do not trust results from regressions that have so many observations and so few covariates. We feel there is a strong likelihood that these regressions will reflect *spurious correlations* due classic omitted variable bias. As a result, we have adopted as an initial working hypothesis that r does not enter as a significant shifter of the distribution $f(a|x,r,c)$, and thus we conclude that the key impact of r on the demand for credit is its effect on customers' propensity to pay for a purchase via installment credit.

4.6 Model Fit

We now discuss the fit of the model. Figures 27, 28, and 29 summarize the ability of the structural model to fit the credit card data. Of course the predominant choice by consumers is to pay their credit card purchases in full by the next installment date: this is the choice made in 93.57% of the customer/purchase transactions in our data set. When we simulate the estimated model of installment choice, taking the x and purchase amounts a as given for the 167,946 observations in our data set, we obtain a predicted (simulated) choice of paying in full at the next statement (i.e. to choose $d = 1$) of 93.56% (this is an average over 10 independent simulations of the model).

Of more interest is to judge the extent to which our model can predict the installment choices made by the customers in our sample, i.e. to predict the incidence of choices $d > 1$. Figure 27 plots the predicted versus actual set of *all* installment choices made the customers in our sample. We see that the model provides a nearly perfect fit of actual installment choices. Figure 28 compares the actual versus predicted choices for the subsample of individuals (both simulated and actual) who chose positive interest installments. We see that once again, the model predicts the outcome we observe nearly perfectly.

The model does slightly overpredict the number of free installments chosen for durations of $d = 2$ installments, and underpredicts the number of $d = 3$ month installments chosen, but only slightly. Overall,

Figure 27: Predicted versus Actual Installment Choices, All Installment Transactions

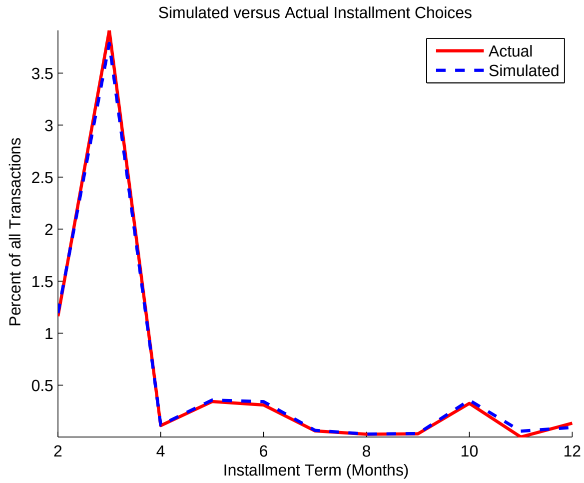


Figure 28: Predicted versus Actual Installment Choices, Positive Interest Installment Transactions

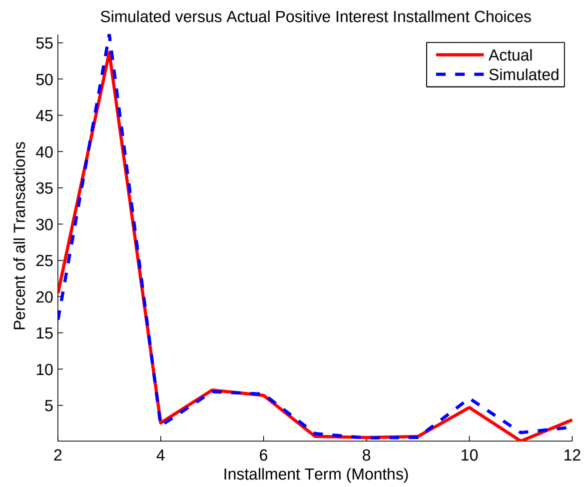
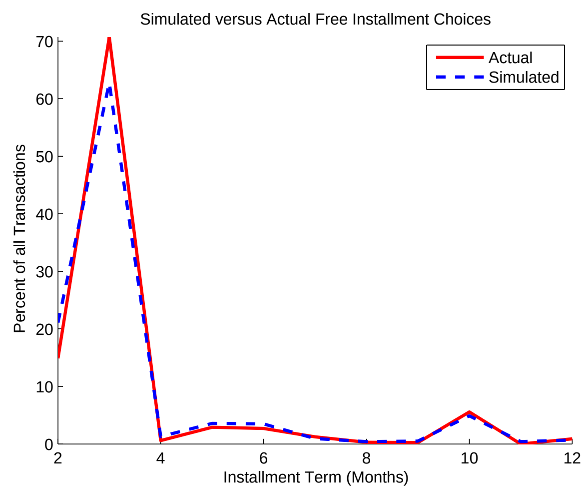


Figure 29: Predicted versus Actual Free Installment Choices



we feel that the model does an excellent job of capturing the key features that we observe in our credit card data. In particular, when we use the simulated data to recreate analogs of the figures presented in section 2, we find that the model succeeds in capturing all of the key features that we observe in the actual data.

We also conducted a battery of Chi-squared goodness of fit tests using the random-cell Chi-squared test of Andrews [1988]. These tests are based on partitioning the dependent variables as well as the covariates entering the model into various "cells" and computing a quadratic form in the difference between the model's predicted probabilities of the customer's choices in the various cells in the partition to the actual frequency distribution of choices in each of the cells. The degrees of freedom depends on the number of cells in the partition less the number of estimated parameters in the model. There are countless ways to partition the space $D \times A \times X \times Z$ where $D = \{1, \ldots, 12\}$ is the choice set, A is the set of (normalized) purchase amounts, X is the set of observed characteristics of customers and Z is a set of all possible merchant code and time dummies that entered the model to predict the probability of a free installment offer. For example, we could partition choices by purchases at various sets of merchants, or over various intervals of time, or on a partition of the amounts purchased (e.g. large transaction amounts versus small tranaction amounts) and so forth. We have done this for many different choices of partitions and while particular values of the Chi-squared statistics are sensitive to how we choose these partitions, we found that with few exceptions the Chi-squared test was unable to reject our model at conventional levels of significance. Given the length of the paper, we decided to omit presentation of the actual test statistics and the correspondence marginal significance values, but we are happy to provide this information upon request.

As we noted in the introduction and elsewhere, our simulations also predict something that we could not otherwise learn from our data without having a structural model: the model predicts that in 17% of 167,946 simulated customer-purchase transactions, the company offers customers free installment opportunities. This estimate strikes us as quite reasonable since if you recall from figure 20 of section 2, the most installment prone "addicts" with *installshare* values greater than 80% were were doing roughly 17% of all of their purchases as free installments. If we assume that the most installment-prone individuals would not pass up many opportunities to purchase items under free installment offers, then this reasoning suggests that our estimated average rate of free installment offers is quite reasonable.

4.7 Model Implications and Counterfactual Simulations

We conclude this section by providing some illustrative simulations of the model and calculating some counterfactual quantities to provide further insight into the model and into the behavior of the individuals in our sample — at least to the extent that the reader trusts that our model provides a good representation of choices consumers actually make.

Figures 30 and 31 illustrate the predicted installment borrowing behavior for two different individuals who are not offered free installment opportunities and so must borrow at an a positive interest rates. In figure 30 we illustrate an “installment avoider” who has an *installshare* of 0, and in figure 31 we illustrate an “installment addict” who has an *installshare* of 83.27%. The credit score happens to be the same for both individuals, equal to 3 (which is a reasonably good score recalling that a score of 1 is the best possible), a moderate installment balance of $ib = 1.85$, and no late payments.

Figure 30 shows that the installment avoider will never choose an installment term of more than three months, and it takes extraordinarily large purchases to motivate this customer to undertake any installment transactions. Even for purchases as large as 10 times the size of the customer’s average statement balance, there is still a 30% chance that this customer will choose $d = 1$, i.e. to pay the purchased amount in full at the next statement date. Figure 31 shows that the installment addict is willing to select installment loans of duration $d = 12$ and this customer’s choice probabilities are much more sensitive to the size of the purchase amount. For small purchases, 20% of the size of this customer’s typical statement amount, there is a 70% chance the customer will choose to pay in full at the next statement, $d = 1$, but a 30% chance of choosing some form of installment loan, with the choice $d = 3$ being the most likely alternative. However when the purchase amount equals the average statement amount for this customer, then there is less than a 10% chance this customer would choose $d = 1$, and the most likely installment terms the customer would choose would be either $d = 3$, $d = 6$, $d = 10$, or $d = 12$. For a purchase equal to 4 times the average statement amount, the chance this customer will select a 12 installment loan is over 60%, with the next most likely alternatives being $d = 10$ and $d = 6$.

Figures 32 and 33 illustrate how the choice probabilities of these two customers are affected when they are given a 10 month free installment offer. Although the free installment offer has little effect on the installment avoider for sufficiently small transactions (e.g. $a = 0.2$), the choice probabilities are dramatically affected by the existence of the free installment option for larger purchase amounts, particularly for the installment avoider. This person had virtually no chance of choosing any installment duration greater than

Figure 30: "Installment avoider" (*installshare=0*)

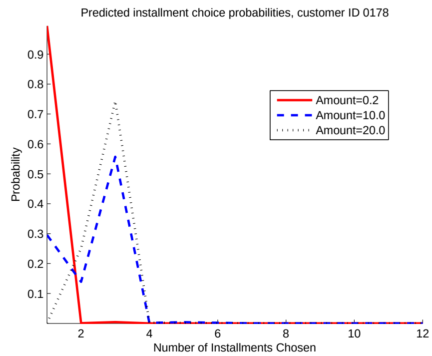


Figure 31: "Installment addict" (*installshare=0.83*)

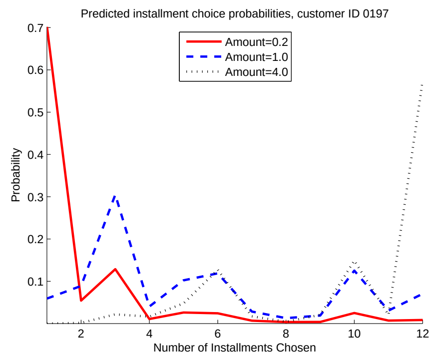


Figure 32: "Installment avoider" (*installshare=0*) with a 10 month free installment offer

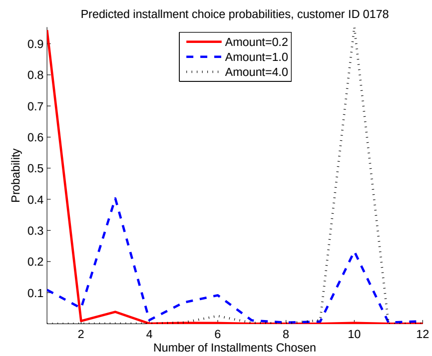
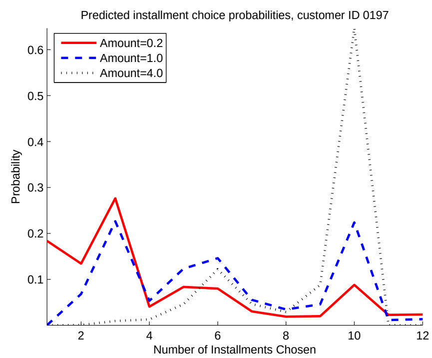


Figure 33: "Installment addict" (*installshare=0.83*) with a 10 month free installment offer



$d = 3$ when facing positive interest rates, however once a 10 month free installment offer is on the table, the customer's chance of taking the 10 month free installment offer starts to increase significantly with the size of the purchase amount a. When $a = 0.2$, the free installment option has very little effect on this consumer's choice probabilities. However when $a = 1.0$ the probability of choosing alternatives $d = 1$ and $d = 3$ fall significantly relative to the case where a free installment offer is not available, and the probabilities of choosing installment durations $d = 6$ and $d = 10$ increase significantly. For even larger purchases, such as $a = 4.0$, the probability of taking the full 10 month free installment offer rises to virtually 100%.

The story is similar for the installment addict, except that this person is motivated to take advantage of the free installment option at lower purchase amounts than we predict for the installment avoider. For a purchase of size $a = 0.2$, the probability of alternative $d = 1$ is only 20% when a 10 month free installment offer is present, compared to nearly 70% otherwise. It is interesting to note that the installment addict is less likely to choose the full 10 month duration of the free installment opportunity than the installment avoider.

This brings us to another key finding: *our model clear predicts that there is a significant probability that customers who choose a free installment will choose a term that is less than the maximum duration offered.* In figures 32 and 33 we see this clearly. For example the blue dashed line in figure 32 shows that if an installment avoider who is purchasing an item that equals the average size of his credit card statement, $a = 1.0$, is offered a free installment with a maximum duration of 10 months, the probability this person will actually choose the free installment at the maximum duration offered, $d = 10$, is less than 25%. Similarly, the solid red line in figure 33 shows that if an installment addict who is purchasing an item of amount $a = 0.2$ and is offered a free installment offer with a 10 month maximum duration, the probability the person will choose $d = 10$ is about 10%.

As we noted in the introduction, simulations of our model for our full sample leads to the prediction that 88% of individuals who were offered (and chose) a 10 month free installment offer also pre-commited at the time of purchase to pay the balance in *fewer* than 10 installments. This pre-commitment behavior, along with the fairly low probability that free installment offers are predicted to be chosen, constitutes what we have termed "the free installment puzzle." Although our econometric model enables us to show this puzzling behavior exists, the model is incapable of explaining *why* individuals in our sample are relatively reluctant to take (or fully exploit) free installment offers. Although we speculated that individuals might have some sort of stigma or fear about some hidden catch or cost associated with taking free installment offers, we simply do not have enough information to be able to isolate the underlying concerns, fears, or other psychological motivations more precisely, or conclude that the behavior is indicative of some form of "time-inconsistent" preferences.

Even though our model predicts puzzling behavior that is inconsistent with standard theories of rational decision making by individuals time-separable discounted utility functions, figures 34 and 35 below show that our model nevertheless does predict downward sloping demand curves for installment credit. These figures present the implied demand curves for the same "installment avoider" and "installment addict" whose choice probabilities we illustrated above. These curves were calculated using the formula for the conditional demand curve for installment credit given by

$$ED(r,x|c) = \left[\int_0^\infty a[1-P(1|a,r,x,c)]f(a|x,r,c)da\right] \tag{28}$$

where $f(a|x,r,c)$ is the customer-specific log-normal distribution for the (relative) amount purchased on any given purchase occasion, conditional on the consumer's decision to use the company's credit card c to pay for the transaction. Note that from our empirical findings in section 4.6, we have no solid evidence that r affects the distribution of purchase amounts, so in calculating these demand curves we simply used customer-specific log-normal distributions $f(a|x,c)$ estimated by maximum likelihood but without including r as an explanatory variable since we found that it does not have any statistically significant effect on a once we included time dummies in the model to control for macroeconomic shocks on spending.

Figure 34 shows that the demand for installment credit by the "installment avoider" is indeed negligible: regardless of the possible credit score, the demand for installment is only a fraction of 1 percent of the average amount of the customer's credit card statement balance. The "installment addict" on the other hand, does have a significant demand for installment credit amounting to approximately an order of magnitude greater than the installment avoider, in relative terms. Thus, depending on this person's credit score, the demand for installment credit in a typical purchase transaction could be anywhere from 10 to 17 percent of the average amount of this person's typical credit card statement amount.

We calculated the demand elasticities for these two customers at the average installment interest rate, 15%, and found in both cases their demand for credit is quite inelastic. The calculated elasticity for the installment addict is -0.074 whereas the demand elasticity of the installment avoider is -0.11. Thus, perhaps not surprisingly the installment avoider has a more elastic demand function than the installment addict, but the important point is both of them have highly inelastic demand curves for credit. This is true for virtually all of the individuals in our sample. Figure 36 plots the distribution of estimated demand elasticities for 607

Figure 34: Estimated installment demand curves for an "installment avoider" (*installshare=0*)

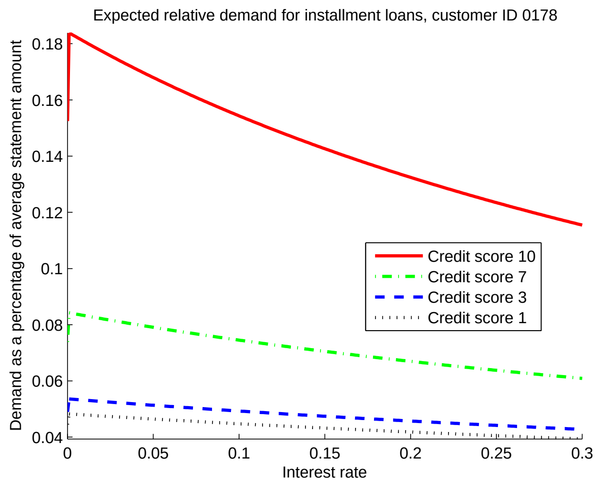


Figure 35: Estimated installment demand curves for an "installment addict" (*installshare=0.83*)

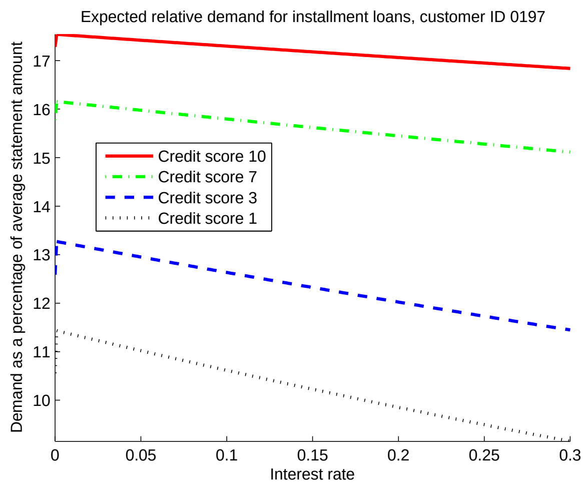
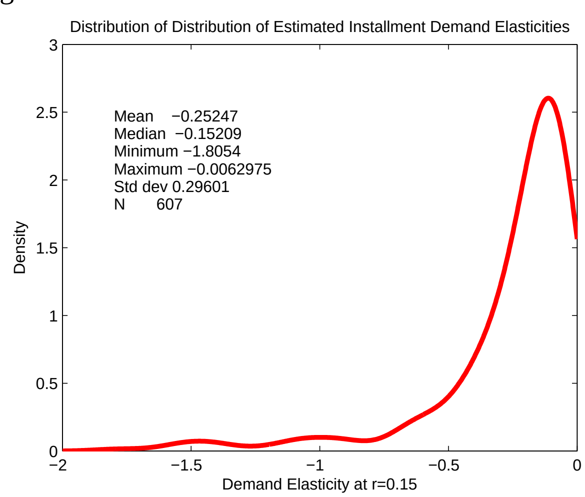


Figure 36: Distribution of Estimated Demand Elasticities



individuals in our sample for whom we had enough data on purchases to calculate reasonable estimates of demand elasticities. We see a very skewed distribution with the lower tail containing a minority of individuals who have relatively elastic demand functions, but the vast majority of individuals have demand elasticities that are quite inelastic and concentrated near 0.

We conclude by examining the optimality of the credit card company's interest rate schedule in light of what we have learned about the demand for installment credit for this sample of customers. Although admittedly, there are hazards to doing an investigation since we do not have a complete model of the demand for credit (in particular, we do not know how interest rates affects customers' decisions about which credit card to use to pay for any given transaction, or how they might affect the total number of shopping trips that the customer might make, i.e. we don't have the data necessary to estimate the functions $\pi(c|x,r)$ and $EN(x,r)$ in the demand curve given in equation (1) of section 3), we argue that such a calculation is reasonable provided we constrain our search for alternative installlment interest rate schedules to guarantee that the customers' expected welfare is no lower under an alternative hypothetical interest rate than the expect under the *status quo.* That is, we solve the following problem

$$\max_{r_2,\dots,r_{12}} \int_0^\infty \sum_{d=2}^{12} [c(a,r_d,d) - c(a,\bar{r},d)] P_+(d|a,x,r_2,\dots,r_{12}) f(a|x) da \tag{29}$$

subject to:

$$\int_0^\infty \log\left(\sum_{d=1}^{12} \exp\{v(d,x,a,r_d)/\sigma)\}\right) f(a|x) da \geq \int_0^\infty \log\left(\sum_{d=1}^{12} \exp\{v(d,x,a,r(x,d))/\sigma)\}\right) f(a|x) da, \tag{30}$$

where $\bar{r}$ is the credit card company's opportunity cost of capital (i.e. the rate at which it can borrow) and $r(x,d)$ is the company's *status quo* interest schedule from equation (7) that we plotted in figure 23 above. The choice probability $P_+(d|a,x,r_2,\dots,r_{12})$ is our model's prediction of the probability that this customer would choose an installment loan of duration d when confronted with a hypothetical alternative interest rate schedule $(r_2,\dots,r_{12})$. The constraint in inequality (30) simply states that the expected utility that the consumer expects from any alternative hypothetical interest rate schedule that the company might offer must be at least as high as the customer expects to receive under the *status quo* schedule. While a fuller specification of the profit maximization problem for the company would probably relax this constraint and instead calculate overall company profits as a sum over all of its customers, accounting for the fact that raising interest rates too much for some customers might cause them to switch to other credit cards or close their accounts entirely, we feel that the constrained optimization problem (29) (30) does give us

Figure 37: Optimal versus *status quo* interest schedules for the "installment avoider" (*installshare=0*)

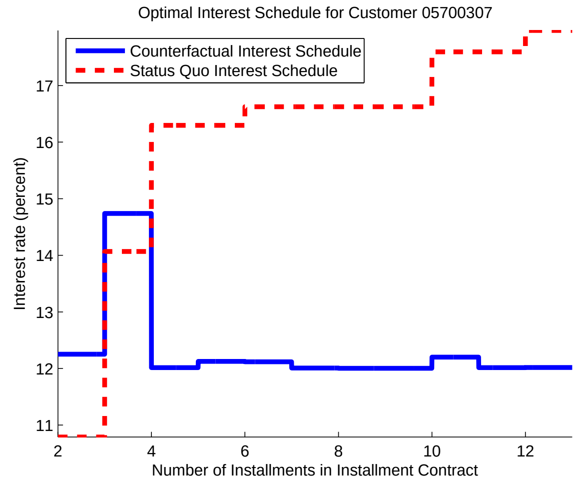
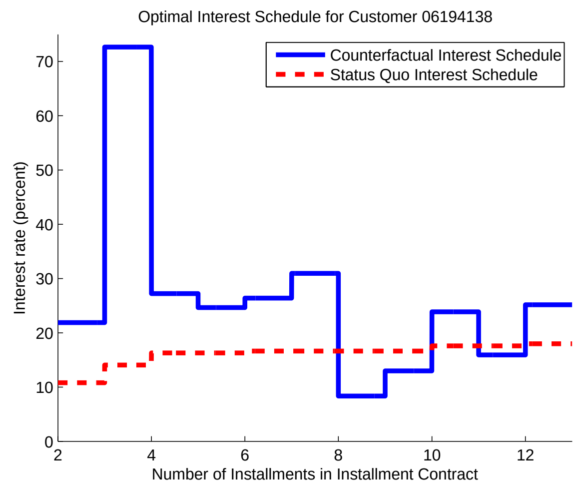


Figure 38: Optimal versus *status quo* interest schedules for the "installment addict" (*installshare=0.83*)



insight whether the company's interest schedule is at least optimal in a *second best* sense. After all, if we can find ways to increase company profits by changing interest rates to its customers without changing the expected welfare they expect from access to the installment borrowing opportunity, the company cannot be maximizing profits in a global sense, since by holding customer welfare constant, we have controlled for the effect of the proposed change in interest rates on the overall demand for and use of the company's credit card by its customers.

Figures 37 and 38 present the optimal schedules that we calculated for the same two individuals that we have studied in our other counterfactual calculations above. These are *customer-specific* interest rate schedules $(r_2,\dots,r_{12})$ that increase the profits the company can expect to receive from these consumers while keeping both customers as well off in an expected utility sense as they are under the company's *status quo* increasing interest rate schedule. Since the company's interest rate schedules are already customer-specific, we believe it is feasible for the company to engage in *first degree price discrimination* and set alternative customer-specific schedules such as the ones suggested in figures 37 and 38.

From figure 37 we see that for the installment avoider, our model predicts the company could increase its profits by generally *lowering* its interest rates except for installment loans with $d = 2$ and $d = 3$ installments, for which its is optimal to increase these interest rates somewhat. The overall decline in interest rates keeps the welfare of this customer unchanged, while enabling the credit card company to extract more surplus from this customer over the durations that the customer is most likely to choose under the relatively infrequent occasions when the customer does do installment borrowing. Note that due to the low rate of use of installments by this customer, overall profits are very low, and even under the alternative interest rate schedule the profits the company can expect from installment loans from this customer are negligible, even though our alternative schedule does increase these (negligible) profits by 10%.

Figure 38 shows a more interesting case, the optimal schedule for the installment addict. Notice that in this case, the optimal interest rate schedule is generally *higher* than the *status quo* interest rate schedule, though the counterfactual schedule is lower at installment loan durations $d = 8$, $d = 9$ and $d = 11$, and the decreases in the rates at these durations are just enough to keep this consumer indifferent between this alternative interest schedule and the *status quo.* In this case, the higher rate of use of installment credit by this customer implies significantly higher profits for the credit card company relative to what it expects to earn from the installment avoider. We calculated profits under the *status quo,* as a fraction of the customer's average credit card statement amount, of 0.5 percent. By adopting the alternative interest schedule in figure 44, we predict that the company can increase its expected profits by over 60% to 0.9 percent of the average statement amount for this customer *per transaction.*

5 Conclusions

The main contribution of our paper is to introduce a new data set on credit card spending and payment decisions, and to study at a high level of micro detail the use of installment transactions, a topic that has not been well studied in previous theoretical and empirical work in economics. We showed that the nature of the installment purchase contract is such that it requires consumers to make individual "micro borrowing decisions" on a *transaction by transaction basis.* Even though the number of consumers in our data set is not huge (fewer than 1000), the panel nature of our data set combined with the frequent use of credit cards by many of the individuals in our sample yield a huge (by economic standards) data set with over 180,000 of these micro-borrowing decisions.

The objective of our analysis was to use this unique set of data to infer customers' demand for credit, since our data also enabled us to identify the *customer-specific* interest rate schedules that the credit card company charges. Unfortunately, due to endogeneity in the setting of customer-specific interest rate schedules (i.e. consumers with worse credit scores who often have the highest need and demand for credit also are assigned the highest interest rates), we found that the traditional "reduced form" econometric methods produced non-sensical estimates of the demand for credit that are *upward sloping* functions of the interest rate r. We found that the use of instrumental variables did not solve the problem since the credible instruments at our disposal (e.g. the CD rate and other measures of the credit card company's cost of credit) are extremely *weak instruments* that do not succeed in producing in downward sloping estimated demand curves for credit.

In order to obtain more credible estimates of the demand for credit we exploited a novel feature of our data: *the company's frequent use of free installment offers.* We argued that the quasi-random way in which these offers are made to the company's customers makes them extremely useful "instruments" an approach that treats free installments as *quasi random experiments* that create extra variation that is helpful in identifying the slope of the demand for credit. Unfortunately, we showed that other standard econometric methods that are designed to exploit such quasi random variation such as *matching estimators* were not adequate, as the estimated treatment effects can easily be misinterpreted as also implying an upward sloping estimated demand for installment credit.

In response to these problems we introduced a flexible discrete choice model of the decision to purchase under installment credit. At each purchase occasion, the customer is modeled as choosing one of twelve installment alternatives, whether to pay the purchased amount in full at the customer's next credit card statement, $d = 1$ (an option that carries a default interest rate of zero), or to purchase the item under installment credit payable in d installments where $d \in \{2,\dots,12\}$ at a positive interest rate that is customer-specific. We accounted for the free installment opportunity as a modification to the customer's choice set: a customer who is given the chance to take out a free installment loan of maximum duration δ may choose from the set $\{2,\dots,\delta\}$ of *free interest options* or can choose to either pay in full, $d = 1$, or borrow for an even longer term $d \in \{\delta+1,\ldots,12\}$ at a positive interest rate. We modeled the choice probability as arriving from a simple cost-benefit tradeoff, where the customer experiences a benefit which we refer to as an *option value function* $ov(a,x,d) = a\rho(x,d)$ that reflects the benefit of the extra flexibility of being able to pay the purchased amount a over d installments.

Offsetting this benefit is a *cost of credit* $c(a,r,d) \simeq ard(30/365)$ and additionally, we assumed that the customer might incur additional *fixed costs* $\lambda(x,d)$ in deciding among the various installment options at check-out time. We showed that the underlying functions ρ and λ can be flexibly specified so that our model can be consistent with a wide variety of rational and more "behavioral" theories of consumer choice. In particular, our model results in a downward sloping demand for credit, even though for certain parameter values our model can predict that consumers should *always* take free installment opportunities when they are offered (and for the maximum duration offered), whereas for other parameter values our model can predict that customers are quite averse to installment borrowing in general and would be even willing to pass up many free installment offers.

We showed that it is possible to solve a major econometric challenge confronting the estimation of our model: namely, that our credit card data are heavily *censored* in the sense that we only observe free installment offers when consumers actually choose them, but the company has no record of other purchase situations where a customer is offered a free installment but did not choose it. Even though it would seem impossible to separately identify the probability of being offered a free installment from the probability of choosing it, we showed that we can indeed separately identify these probabilities. What we found was surprising: even though only 2.7% of the transactions in our data set were done as free installments, our model predicts that consumers face free installment offers in approximately 20% of all the transactions they make.

The *free installment puzzle* results from this key finding, namely that customers in our data set are predicted to frequently pass up "free" borrowing opportunities. Further, we also showed that in the minority of cases (15%) where customers did choose the free installment offer, there was a very high probability (approximately 88% for a 10 month free installment offer) that the consumer would pre-commit to a choice of a loan duration that is *shorter* than the maximum duration allowed under the offer. These decisions present a challenge to traditional economic models of rational, time-separable discounted utility maximization. Pre-commiting to "suboptimal" choices can be evidence that individuals have more complicated *time inconsistent* preferences for which this type of pre-commitment can be welfare improving by constraining future options and the potential "temptations" that current borrowing poses for their welfare of their "future selves."

While our model does raise new puzzles, it also resolves others. For example, even though our model generates downward sloping demand for installment credit, it is nevertheless consistent with the counterintuitive estimated treatment effects from the matching estimators that we presented in section 2. The matching estimators predict that consumers spend significantly *less* per free installment transaction than they do for positive interest installment loans, a finding that is easily misinterpreted as a prediction that customers have upward sloping demand for credit. Our model predicts that free installment offers significantly lower the threshold at which consumers are willing to make an installment purchase, thereby lowering the average size of a free installment transaction. But since this lower threshold also implies a greater number of transactions will be done via installment, our model predicts that free installments increase *total installment borrowing* even though the average size of a free installment purchase is lower.

While we believe we have provided credible evidence that this type of pre-commitment behavior is common (something that few other non-experimental empirical studies have done so far, to the best of our knowledge) we still refer to our findings as the "free installment puzzle" since our data are not rich enough to delve deeper into the psychological rationale for these decisions. Besides time-inconsistent preference explanations, there are other potential "behavioral" explanations for these choices, including social stigma against the use of installment credit and the scarring effect of past overuse of installment credit. Since installment credit decisions are made at the check out counter in a public setting, the potential stigmatization effect cannot be discounted (similar to the way the use of food stamps at check out counters may be a source of embarrassment for consumers in the U.S.). We believe a distinct possibility is that our findings reflect the chastising effects of the rapid growth and sudden bursting of a large "credit card bubble" in the country just prior to the period of our data, and that this experience could have significant scarring effects that made many consumers hesitant to take advantage of installment credit opportunities given that excessive use of installment credit had created so many problems for this country in the very recent past.

While we presented calculations that suggest that the credit card company's interest rate schedule may not be optimal, we cannot provide any definite conclusions whether the company's use of free installments is an effective policy or not. We did show that the people who are among most likely to respond to free installment offers — individuals with high values of the *installshare* variable — also tend to have worse creditscores but also tend to be more profitable customers. Although the response to free installment offers seems small even for individuals with high values of *installshare* our analysis is unable to address the question of whether the primary effect of free installment occurs if customers switch credit cards at the checkout counter in order to take advantage of free installment offer provided by one credit card but not another.

This point is connected to our final point, namely that an important limitation of our study is that our data only allows us to study credit decisions for customers of a single credit card company. Of course, customers have a choice of many different ways to pay at the check out counter, including using cash or other credit or debit cards. Though we did find that demand for installment credit is generally quite inelastic, it is important to remember that our finding is *conditional on the use of this particular credit card* and thus we have additional problems due to the choice-based nature of our sample of data. In the future, it would be important to study consumer choice over multiple alternative sources of payment similar to the study by Rysman [2007] who studied payment choices across multiple different competing credit cards. It seems reasonable to suppose that the overall demand function for credit will be more elastic when we open up the analysis to consider all of the possible alternative means of payment.

References

Sule Alan, Ruxandra Dumitrescu, and Gyongyi Loranth. Subprime consumer credit demand – evidence from a lender's pricing experiment, 2011.

Donald W. K. Andrews. Chi-square diagnostic tests for econometric models: Theory. *Econometrica*, 56-6: 1419–1453, 1988.

Nava Ashraf, Dean Karlan, and Wesley Yin. Tying odysseus to the mast: Evidence from a commitment savings product in the phillipines. *Quarterly Journal of Economics*, 121-2:635–672, 2006.

Lawrence Ausubel and Haiyan Shui. Time inconsistency in the credit card market, 2005.

Maro Casari. Pre-committment and flexibility in a time decision experiment. *Journal of Risk and Uncertainty*, 38:117–141, 2009.

Faruk Gul and Wolfgang Pesendorfer. Temptation and self-control. *Econometrica*, 69-6:1403–1435, 2001.

Marc Henry, Yuichi Kitamura, and Bernard Salanié. Identifying finite mixture models, 2011.

David Laibson. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112:443–477, 1997.

Marc Rysman. Empirical analysis of payment card usage. *Journal of Industrial Economics*, 60:1–36, 2007.

R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23: 163–180, 1955.