Consumer Credit: Learning Your Customer's Default Risk from What (S)he Buys

Annette Vissing-Jorgensen
Haas School of Business, University of California Berkeley, NBER and CEPR

June 2, 2021

Abstract

The use of large account-level data sets on spending patterns is part of the big data revolution in household credit scoring in the finance industry. However, due to limited data access, academics have little insight into which predictors are successful and what the economics of their predictive power is. Based on the insight that detailed consumer surveys contain similar information to that in account-level data from debit and credit cards, I use the US Consumer Expenditure Survey (CEX) to study the link between spending patterns and consumer credit outcomes (paying positive finance/interest/late charges on consumer credit). I supplement this with analysis of a large account-level data set from a Mexican retail chain that sells durables on credit and for which I can measure loan default. In both datasets, I find that spending on categories related to entertainment (such as video, audio, magazines, newspapers, toys and pets) predicts worse credit outcomes. To understand the underlying economics, I conjecture that impatient households are more likely to spend on particular categories. I use smoking and lower education as proxies for impatience in the US CEX data and show that spending categories that predict paying positive finance/interest/late charges on consumer credit also tend to predict smoking and lower education. Cross-sectional heterogeneity in patience thus appears central for understanding the cross-section of consumer credit outcomes.

I thank the company that provided the Mexican data set and Jose Guillermo Diaz for research assistance. I also thank James Choi, Michael Haliassos, Justine Hastings, Gregor Matvos, Jesper Rangvid, and seminar participants at the University of California Berkeley, New York University, Kellogg School of Management, Johns Hopkins University, Notre Dame, Illinois Institute of Technology, Board of Governors of the Federal Reserve, the NBER Behavioral Finance Meeting, the NBER Summer Institute (ACR), the American Economic Association Meeting, European Finance Association Meeting, the Conference on Consumer Credit and Bankruptcy in Cambridge UK, CEPR ESSFM (Gerzensee), and the CFS Conference on Household Finance Greece for their comments.

1. Introduction

The field of household finance initially focused mainly on the asset side of households' balance sheet, studying wealth accumulation and portfolio choice. This changed with the financial crisis, which initiated a large literature on household mortgage debt, thus shifting some of the field's attention to household liabilities. Aside from mortgages, the main household liability is consumer credit, consisting of credit card debt, auto loans and student loans. As of the end of 2020, consumer credit in the US stood at $4.2T, compared to mortgage debt of $11.2T (US Financial Accounts). Despite its substantial size, consumer credit remains much less studied. Reflecting the state of the field, in their excellent 121-page survey of household finance, Guiso and Sodini (2013) devote only a couple of pages in their final section to consumer credit.

New developments suggest that increased focus on consumer credit is warranted. In particular, a large part of the big data revolution in the finance industry focuses on modeling consumer credit risk. Most famously, Ant Group (an affiliate of Alibaba Group) has $155B of credit outstanding to Chinese consumers, issued by leveraging information on their spending habits and payment histories.[1] In the US, much of the innovation in fintech has similarly focused on disrupting the markets for household borrowing, including consumer credit. A second recent development is the sharp increase in the student loan component of consumer credit, which has led to a heated debate about the fairness of education funding as well as the ability of borrowers to repay.

In this paper, I focus on the first of these trends, the use of large and novel data sets for credit risk assessment. Novel data used to model consumer credit risk include granular transactions data from credit or debit cards, bill payment data (rent, utilities), social media data, mobile payment data, and clickstream data (the digital footprint created from using a web site).[2] Since the novel information sources are not ones academics generally have access to, there is a disconnect between the explosion of big data efforts in industry and our (academia's and the general public's) understanding of the economics of why such analysis works. A pioneering paper showing the value of digital footprint for credit scoring is Berg, Burg, Gombovic and Puri (2020). Using data from a German online vendor, they document the value for default prediction of variables obtained as part

[1] See Reuters (2021), https://www.reuters.com/article/us-china-ant-group-consumer-finance/chinas-ant-to-boost-consumer-finance-unit-capital-as-it-restructures-micro-lending-sources-idUSKBN2AQ11C.
[2] See, e.g., World Bank (2019) and FICO's blog at https://www.fico.com/blogs/using-alternative-data-credit-risk-modelling.

of the online shopping process, such as the operating system used (iOS or Android, a proxy for cellphone cost and thus income) and whether customers have their names in their e-mail address (capturing an aspect of their personality). I seek to understand the value of transactions data on consumer spending mix (what people buy, as opposed to how they buy it) for understanding credit risk. In particular, I am interested in whether the type of good or service purchased has predictive power for understanding credit outcomes and what this can teach us about the fundamental drivers of household borrowing and credit risk.

To study what one can learn from account-level data, I take two approaches. The first approach is to observe that high-quality household surveys contain much of the same information as is available in detailed account-level data. For example, a bank with access to your checking account, debit card, and credit card transactions could observe your spending patterns in much the same way as it is obtained in survey data from Consumer Expenditure Survey (CEX). Using data for 64,000 consumers with consumer credit (excluding mortgages and vehicle finance) in the Consumer Expenditure Survey over the 1988Q1-2013Q1 period, I document how consumer credit interest and finance charges relate to spending patterns. The main novel finding from this analysis is that spending on a series of spending categories related to entertainment is significantly positively related to paying finance/interest/late charges. These categories include several categories related to video and audio (purchases and rental of video cassettes/tapes/discs/films, video game hardware/software, records/CDs/audio tape purchases and mail order purchases), categories related to reading (including non-subscription magazines and newspapers), categories related to pets (pet purchase/supplies/medicine and vet costs), a category for toys/games/arts/crafts/tricycles, and a category for lotteries. I also find that information about household borrowing of other types (mortgages and vehicle finance) is correlated with finance/interest/late charges on consumer credit. This confirms the value of traditional credit scoring which predicts creditworthiness based on past borrowing and repayment behavior.[3]

[3] For example, in the calculation of an individual's FICO score (a commonly used measure of credit risk in the U.S.), 35% weight is given to on time payment of past debt, 30% to the current amount of debt of various types, how many accounts the individual has, and how large the debt is relative to the total available credit, 15% to the length of time of credit history, 10% to the number of new accounts and recent requests for credit, while 10% is given to the mix of credit (credit cards, installment loans, finance company loans, and mortgages) used in the past (Fair Isaac Corporation (2005)). For an example of the use of FICO scores to predict default see Keys, Mukherjee, Seru and Vig (2010) in the context of mortgage delinquencies.

However, this type of information is generally reported by lenders to credit bureaus and thus does not necessitate study of account-level data. Furthermore, from the perspective of understanding the underlying economic drivers of debt and default, predicting default based on past repayment behavior is not informative.

My second approach to understanding the value of account-level data for credit scoring is to study default in a large account-level dataset from one of the largest retail chains in Mexico. The chain sells many types of durables on credit. This allows me to study default, not just the paying of interest or finance charges. While paying interest or finance charges is very likely to have predictive power for default, actual data on default allows direct prediction of lender losses. The Mexican data set is large both in terms of the number of borrowers covered -- about half a million -- and in its panel dimension, with monthly data available from January 2005 to August 2009. A unique feature of the data, which enables me to directly study the link between which goods are purchased and subsequent loan losses for the lender, is that each good purchased at this retailer has its own loan associated with it. For example, a customer may first take out a loan to buy a washing machine and then later take out a separate loan to buy new tires for her car.

Results from the Mexican data are consistent with the survey-based results: Loans given for purchase of products that provide various type of entertainment have much higher default rates than loans given to fund purchases of less exciting items. Specifically, lender loss rates (measured as the ratio of the amount not repaid to the size of the loan) are about 21% on loans for electronics such as cell phones, stereos, and TVs but much lower (below 12%) on loans for appliances, kitchen equipment and furniture. These differences in default losses across product categories are robust to controlling for characteristics of the loan contract (e.g. the size of the loan and interest rate on the loan), demographics, and more standard measures of credit risk based on past repayment behavior, and the differences do not diminish substantially with how long the borrower has been a customer. From a statistical perspective, this implies that which products people borrow money to buy is a useful additional predictor of subsequent default, above and beyond known predictors documented in past work.

Exploiting the panel dimension of the Mexican data, I estimate models of loss rates that include customer fixed effects. With these fixed effects included, differences in loss rates across product categories are economically small. This indicates that differential loss rates across product categories are driven mainly by which types of individuals buy particular products, as opposed to being product-specific features. In other words, customers who tend to buy electronics generate high lender loss rates both when they buy electronics and when they buy other products.

To understand why entertainment-related spending relates to paying finance/interest/late charges on consumer credit in the US data and to default in the Mexican data, I study whether there is a link between spending patterns and impatience. In standard models of household life-cycle consumption and savings, consumers borrow to smooth consumption across age (given a mismatch in timing between income profile and desired spending, and including to smooth consumption of durables) and across states of the world (smoothing transitory shocks to income or to needed expenditures). In this framework, repayment difficulties could be driven by:

- Bad planning: A household took on more debt than it would if it fully understood the optimization problem upfront. One could include (naïve) time inconsistent preferences in this category.
- Sufficiently bad news about realized income and expenditure needs.
- Impatience: For given expectations about income processes and possible expenditure shocks, impatient households will be willing to borrow more upfront even if this carries a greater risk of default and low consumption later.

Using the CEX data, I consider two proxies for impatience: Smoking (which can be identified by positive cigarette and other tobacco spending) and fewer years of education. Smoking provides immediate rewards at the cost of later health problems while completing education requires sacrificing current income and leisure for increased income later. I show that the same spending categories that predict paying positive finance/interest/late charges on consumer credit also tend to predict smoking and lower education. This suggests that impatient households are more likely to spend on particular categories and thus that cross-sectional heterogeneity in patience is central for understanding the cross-section of consumer credit outcomes.

Understanding what drives heterogeneity in borrowing and default behavior across households in the market for consumer credit is likely to be informative for other loan markets and for understanding consumption and savings behavior more generally. In particular, if substantial preference heterogeneity can be documented within a sample of households who all have consumer debt, one would expect an even larger degree of preference heterogeneity in the full set of households, with correspondingly broader implications for explaining heterogeneity in household net worth. Consistent with the importance of heterogeneity in time preferences, Calvet, Campbell, Gomes and Sodini (2021) use administrative data from Sweden and a model of saving and portfolio choice to document modest heterogeneity in risk aversion but considerable heterogeneity in the time preference rate and elasticity of intertemporal substitution.

The paper is organized as follows. Section 2 contains the analysis using the US data from the Consumer Expenditure Survey, section 3 analyzes the Mexican data set and section 4 concludes.

2. Consumer credit in the US Consumer Expenditure Survey

2.1 Sample construction

Over the period 1988Q1-2013Q1, the CEX contains information about how much households paid in finance, interest and late charges over the past 12 months on consumer credit.[4] This information is available for households in their fifth and last interview.

The CEX measure of consumer credit includes credit card debt (from major credit cards, store credit cards, or gasoline credit cards), store installment credit, credit from financial institutions (including banks, savings and loans, credit unions, finance companies, and insurance companies), credit from health care providers (doctors, dentists, hospitals, and other medical practitioners) and other credit sources. I exclude credit from health care providers from my analysis since it may be driven by different factors than other consumer credit. The CEX consumer credit measure excludes mortgages, home equity loans, vehicle loans, and business related loans.

I combine the consumer credit information with data on demographics and detailed expenditures. The CEX contains hundreds of spending categories (referred to as UCC codes). I include all categories used by the CEX in its calculation of total expenditure, except for those in the "personal insurance and pensions" category which I view as savings rather than consumption expenditure. There are changes over time to the exact expenditure category used, as some categories are dropped and others added. In total, 619 categories appear across the 1988Q1-2013Q1 period. For each expenditure category I calculate a given household's spending on this category across available

[4] This information is available in the "FNB" files contains in the detailed expenditure files. Information is also available about the amount of consumer credit borrowing using the "FNA" files.

interviews (some households have a fifth interview despite having missed one of the earlier interviews). I drop spending categories for which less than 1% of households report positive spending (across survey quarters in which the category is used). This results in 448 remaining spending categories, listed in Appendix Table 1 for reference. I include quarterly time dummies as controls in all regressions to account for the fact that not all spending categories are used in all survey quarters.

My regression sample uses one observation per household, with values for variables referring to the last interview, except that the household's earlier expenditure data are also used to calculate expenditures for each category for each household. The sample has 157,553 distinct households, of which 66,997 (43%) have positive consumer credit based on the CEX question which asks for balances as of the 1st of the current month. I focus on the set of 63,691 households with positive consumer credit and positive after-tax income. Of these, 69% have positive finance/interest/late charges on consumer credit.[5] Among those with positive finance/interest/late charges, the average is $575 and the median $221 (for a 12-month period). Credit cards account for 87% of observations with positive consumer credit and finance/interest/late charges on credit cards account for 79% of consumer credit finance/interest/late charges.[6]

2.2 Univariate analysis

I start with a univariate approach to link finance/interest/late charges to expenditure patterns. For each spending category *c*, I estimate a linear probability model

$$\text{D(Positive finance/interest/late charges)}_i = \gamma_c * \text{D(Positive category spending)}_{i,c} + x_i'\beta + \varepsilon_i \qquad (1)$$

where *i* denotes household, *c* denotes a spending category, x_i is a vector of household-level controls and data are for 1988Q1-2013Q1.[7] As controls, I include the log and log squared of total expenditure (excluding personal insurance and pensions), log and log squared of after-tax income. For both spending and income, I use real variables, adjusting by the consumer price index. Controls

[5] These percentages are unweighted. Using CEX weights 37% have positive consumer credit and 60% of those with positive credit have positive finance/interest/late charges on consumer credit.

[6] No household is in its fifth interview between April and December 2005 due to sample redesign so this period is dropped in my analysis. A few households have negative total expenditure and are dropped.

[7] The next version will also exploit the intensive margin of both spending and finance/interest/late charges, using Tobit models and modeling spending shares across categories.

also include the respondent's age and age squared, family size, a dummy for the respondent being male, a dummy for the household residing in a rural area, and time dummies (quarterly).

Table 1 reports the results of this analysis. I report the t-statistic on γ_c and the value of γ_c along with the fraction of consumers with positive spending in the category. The table is sorted in decreasing order of the value of the t-statistic on γ. Rather than reporting results for all 448 regressions, I report the 50 categories with the largest t-statistics (in absolute value). I draw the following lessons.

First, even controlling for total spending and income, along with a set of demographics and time dummies, many categories have significant predictive power (in a cross-sectional sense) for paying positive finance/interest/late charges. Many of the category coefficients are substantial, implying a several percentage point higher probability of paying finance/interest/late charges for households with positive spending on the category. This is consistent with the usefulness of account-level data for (potential) lenders. It appears easier to predict high-risk than low-risk borrowers. Of the 50 strongest predictors, 47 have positive coefficients while only 3 have negative coefficients.

Second, consistent with the usefulness of traditional credit scoring some of the strongest predictors of positive finance/interest/late charges on consumer credit are related to payments on other borrowing or banking products (highlighted in light blue). These categories include mortgage interest on owned dwellings, interest on lump sum home equity lines of credit for owned dwellings, automobile finance charges, truck finance charges, and interest paid on other vehicles. This finding is not surprising given that consumer credit tends to be much more expensive than mortgages or vehicle loans, implying that households will tend to take on consumer credit after available lower-cost options are used. Consistent with consumer credit finance/interest/late charges indicating some level of financial distress, the category with the strongest predictive power is positive payments for checking accounts and other bank service charges (this category includes below minimum balance fees).

Third, and most interesting, many of the categories with the statistically strongest predictive power for predicting positive finance/interest/late charges on consumer credit are what one could refer to as entertainment-related items (highlighted in pink). I classify 16 of the top 50 categories as such, all having positive coefficients. These 16 include five categories related to video and audio, five categories related to magazines, newspapers, and books (of which four are single-copy as opposed to subscription), two pet-related categories, toys/games/arts/crafts/tricycles, lotteries, film (for cameras), and telephone service.[8] There is obviously a level of subjectivity in this classification. I intend to capture spending that relates to entertainment of various types, broadly defined.

In terms of categories with negative coefficients, the three strongest predictors are safe deposit box rental, capital improvement costs (labor/material) for owned dwellings, and electricity for owned vacation homes. While these three categories are all luxuries in the sense that the budget share positively related to total consumption, I do not find a strong relation across spending categories between γ_c (or the t-stastitic on γ_c) and necessity/luxuriousness (omitted for brevity). Having a safe deposit box is perhaps indicative of conscientiousness.

2.3 Multivariate analysis

To assess which predictors are important in a multivariate setting, I use a LASSO approach. For a linear model, $y=\beta_1 x_1+ \beta_2 x_2+\ldots+\beta_p x_p+\varepsilon$, the LASSO objective is to minimize

$$\frac{1}{2N}(y - X\beta)'(y - X\beta) + \lambda \sum_{j=1}^{p} |\beta_j| \tag{2}$$

The first term is the OLS objective. The second term penalizes non-zero coefficients. Because of the kink in the absolute value function, some coefficients are set to zero. The value of λ determines how sparse the chosen model will be. There are various methods to choose λ. I pick it using the extended Bayesian information criterion, which tends to result in a smaller set of predictors than other criteria.[9] The advantage of LASSO over OLS is to avoid overfitting in settings with many regressors. It has better statistical properties than simply adding or deleting variables using a stepwise OLS approach.

The multivariate model I use is

$$\text{D(Positive finance/interest/late charges)}_i$$

$$= \gamma_1 \text{*D(Positive category spending)}_{i,1} + \ldots + \gamma_C \text{*D(Positive category spending)}_{i,C} + x_i'\beta + \varepsilon_i \tag{3}$$

[8] The three categories related to single-copy magazines and single-copy newspapers are really two categories. They are combined into one from 2006 onward. Neither exists pre-1994 where subscription and singly-copy magazines are combined into magazines and similarly for newspapers.
[9] I use the STATA package lasso2.

with the same controls and sample as for the univariate approach. I do not impose a penalty on the time dummies included in the controls to ensure that they are included in the model. This is needed to obtain a meaningful model because not all spending categories are used in all time periods.

The result of the LASSO estimation is shown in Table 2. The purpose of the table is to show which variables are selected. The table presents the spending mix variables chosen by LASSO, sorted based on the post-estimation OLS t-statistics. I also show the t-statistic for γ_c and the coefficients for γ_c based on the post-estimation OLS regression (which uses only the chosen variables). These are biased and are not to be used for inference. I include them as indicators of the relative importance of variables. The controls chosen by LASSO are shown at the bottom of the table. LASSO (with the chosen approach to pick λ) selects a total of 38 spending categories, along with log real after-tax income and dummies for the respondent being male and the household living in a rural area. Of the 38 spending categories, 32 have positive γ_c, consistent with the univariate results. Of the 32, 7 relate to other debt and banking and 12 are entertainment-related categories.

To get a sense of the predictive power of various sets of variables, I re-estimate the final model using logit and calculate the area under the ROC curve (called AUC), a standard measure of fit in binary models which lies between 0.5 and 1.[10] Using only the time dummies and the three chosen control variables results in an AUC of 0.605. Adding the 7 variables related to other debt and banking increases the AUC to 0.646. Also adding the 12 entertainment-related categories increases the AUC to 0.653, while adding the remaining 19 spending category variables increases the AUC to 0.666. Iyer, Khwaja, Luttmer and Shue (2016) state that "an AUC of 0.6 or greater is generally considered desirable in information-scarce environments, and AUCs of 0.7 or greater are the goal in more information-rich environments". This is in line with my finding that the AUC increases from 0.605 with just a few variables (and time dummies) and to 0.666 when adding the 38 spending dummies retained by LASSO.

2.4 Mechanism

Where does predictive power of spending mix for consumer credit outcomes come from?

The role for expenditure shocks appears to be modest. Of the categories in Table 1 and 2, few represent expenditure shocks -- physician services, vet services, and prescription drugs perhaps.

[10] For consistency across results, the next version will use logit models throughout.

This is not mechanical. For example, the CEX has many repair categories that could have emerged as important but do not.

The categories related to other borrowing and banking are not informative from an economic mechanism standpoint. If the same people tend to borrow in many ways, that still does not say anything fundamental about why they made this choice and others did not. The predictive power of categories related to entertainment is more informative. Since the regressions control for income, it is not simply the case that those who spend money on entertainment categories also spend time on them and thus have lower income and therefore worse credit outcomes. Instead, consider the possibility that more impatient households are more likely to spend on particular categories. For example, they may put a higher value on immediate experiences and excitement and be willing to borrow to pay for it. The standard discounted utility framework does not link *inter*-temporal preferences to *intra*-temporal preferences. Utility is assumed separable across periods, and optimization is done first across periods and then within period. However, those who are less patient may also have stronger preferences for goods and services that provide experiences. Perhaps they find it particularly utility-reducing to have to delay experiences.

To assess whether spending mix is related to discounting, I need a proxy for discounting available in the CEX data. I consider smoking by someone in the household as one possible proxy of the household's discount rate, and define smoking based on having positive spending on cigarettes or other tobacco products. Smoking is associated with immediate utility but negative health consequences later, so those who value the future less should be more likely to smoke. Consistent with smoking being correlated with impatience and impatience being important for finance/interest/late charges, smoking spending appears as a significant predictor in both Table 1 and 2.[11] For each spending category c, I estimate the same linear probability model as in (1), replacing the dependent variable with a dummy for whether household i has positive spending on smoking:

$$\text{D(Positive smoking spending)}_i = \gamma_c^{smoke} * \text{D(Positive category spending)}_{i,c} + x_i'\beta + \varepsilon_i \qquad (4)$$

I use the same controls and sample. I interpret a positive value for γ_c^{smoke} for a given category as an indication that this category tends to attract less patient households. In the cross-section of 448

[11] In the Survey of Consumer Finances, respondents are asked about both smoking and a survey-based measure of planning horizon. I have confirmed that those with shorter planning horizons are in fact more likely to smoke.

spending categories, I relate the estimated values of γ_c^{smoke} based on estimations of equation (4) to the estimated values of γ_c from estimations of equation (1). If the same categories that predict smoking (i.e., impatience) also predict positive finance/interest/late charges, this would suggest that the predictive power of spending mix for finance/interest/late charges comes from a relation between spending mix and discounting. I assess the relation between γ_c^{smoke} and γ_c based on both their t-statistics and coefficient magnitudes.

Table 3, column (1) shows the results of a cross-sectional regression of the t-statistic for γ_c^{smoke} on the t-statistic for γ_c. I include the fraction of households with positive spending on a given category as a control, since this fraction may affect both sets of t-statistics. The two t-statistics are significantly related suggesting that categories that are statistically more tightly related to smoking are also statistically more tightly related to positive finance/interest/late charges for consumer credit. As shown in column (2), using coefficient magnitudes rather than t-statisics, γ_c^{smoke} and γ_c are also positively related.[12] Figure 1, top left, shows a bin-scatter plot relating the t-statistic for γ_c^{smoke} and the t-statistic for γ_c while Figure 1, top right, shows a bin-scatter plot relating γ_c^{smoke} and γ_c.

As an alternative proxy for patience, I consider years of education of the respondent. Education requires a delay of consumption and study effort upfront, returning higher consumption later. More impatient individuals should thus tend to choose less education.[13] If spending mix is related to finance/interest/late charges because it reflects discounting, then spending on categories with positive γ_c should tend to predict shorter education. To test this idea, for each spending category *c*, I estimate the following relation

$$(\text{Years of education of respondent})_i = \gamma_c^{education} * D(\text{Positive category spending})_{i,c} + x_i'\beta + \varepsilon_i \quad (5)$$

using the same controls and sample as for equation (1) and (4). I then relate the t-statistics for $\gamma_c^{education}$ and the t-statistics for γ_c, and also relate the coefficient magnitudes. The results are shown in column (3) and (4) of Table 3 and illustrated with bin-scatter plots in the bottom two graphs in Figure 1. The negative relations between the two sets of t-statistics (Table 3, column 3) or the two

[12] The gammas and their t-statistics are estimated, which adds noise to the relations in table 3. Accounting for estimation error would likely only strengthen the results.
[13] In the Survey of Consumer Finances, a shorter planning horizon does in fact predict longer education.

sets of estimated coefficients (Table 3, column 4) provide further support to the idea that spending mix predicts consumer credit charges because spending mix reflects patience.

While these results point toward cross-sectional heterogeneity in patience as an important driver of cross-sectional differences in consumer credit outcomes, much more work is needed to understand how impatience drives borrowing and borrowing outcomes. Are those with higher entertainment spending more likely to be hyperbolic discounters? Or, do more impatient household have time consistent preferences with higher discount rates? If so, is that discount rate the same for all types of consumption, or is it higher for certain categories like entertainment?[14] Alternatively, perhaps we should think of impatience as related to impulsiveness and impulse shopping. Gathergood (2012) document a relation between impulse buying and financial distress. A role for impulsiveness would fit with the importance of the spending categories for singly-copy magazines and newspapers, since these are likely to be unplanned purchases. Are consumers who spend a large fraction on the entertainment-related categories succumbing to a desire for immediate gratification both when shopping and when deciding what to do with their time? Adding patience and impulsiveness-related survey questions to the CEX, or combining account-level data with customer surveys at financial institutions would be informative for assessing this idea.[15]

3. Consumer credit at a Mexican retail chain

To study not consumer credit default, a more severe outcome than paying finance/interest/late charges, I turn to a large dataset obtained from a Mexican retail chain.

3.1 The basics of the data set

The data sets consists of information about 499,906 new customers who purchased one or more products on credit at one of the largest Mexican retail chains between January 2005 and December 2006. During this time period, this set of customers made a total of 1,364,864 credit-financed purchases. The payment history of these purchases is followed up to August 2009.

[14] Consider a setting with two periods (1 and 2) and two goods (A and B).two-period, two-good setting. The standard framework expresses lifetime utility as $u(c_1^A, c_1^B) + \beta u(c_2^A, c_2^B)$. Do those with a stronger preference for A simply have different β? Or, is utility characterized by a different β for each good, $u(c_1^A) + \beta^A u(c_2^A) + v(c_1^B) + \beta^B v(c_2^B)$. Think of good B representing entertainment and being associated with a lower β, with some households having a stronger preference for B relative to A.

[15] The field of marketing has developed scales to measure impulsive shopping based a several questions (see Rook and Fisher (1995) or Puri (1996)).

The retail chain which made the data available makes about 90% of its sales on store credit, with the remaining sales paid in cash or using credit or debit cards. The chain was founded several decades ago, is now represented in all 32 Mexican states and has millions of customers. During the 2005-2006 period, the chain was not represented in a few states. The purchases in my sample are made across 220 different stores.

The company's target customers are middle and lower income households. 88% of customers in my sample have monthly household incomes below 16,800 pesos ($1,268). 52% have monthly household incomes below 4,200 pesos ($317). For comparison, in the Mexican population as a whole, 85% have monthly household incomes below 16,800 pesos, while 26% have monthly household incomes below 4,200 pesos (calculated using ENIGH 2005, a national household survey conducted by the Mexican government). A large part of the company's success is attributable to its ability to sell products on credit to this segment of households, many of whom have no other sources of credit. One of every five employees work in credit supervision.

The data set provided by the company contains monthly information of four types. First, the company collects and updates customer demographics, specifically age, gender, marital status, household income, education, home ownership, years at current address, and household size. Second, information is provided about any movements in the customer's accounts. Movements include new purchases made, payments on past purchases, assignment of additional interest (due to late payments) and merchandise returns. The data set covers purchases made on credit only. For each new purchase, the data set contains information about the store at which the purchase took place, the amount of the purchase, the size of the down-payment, the interest rate on the loan, the term of the loan, and the type of product purchased. Third, monthly data are available for the customer's account balances, the customer's track record of repaying loans, and the customer's credit limit. Fourth, the data contains information about ``lost loans", meaning loans on which the company has given up collecting any further payments. For such accounts, records are kept on the date of purchase, the date the account was declared lost, and the amount of the loss to the company. The average time between date of purchase and date the account was declared lost is just over two years. This reflects the fact that loan terms are 12 or 18 months and it takes the company a while to determine whether any further payments can be collected on a given loan. The two year lag between a purchase and the typical date of a loan being declared lost motivates my focus on purchases made in 2005 and 2006. Since the sample runs to August 2009, this provides sufficient panel dimension to follow the outcome of each loan.

Of particular importance for my analysis is the information about the type of good purchased and the way loans are made. For the purchases made in 2005 and 2006 I have for each purchase a basic product description such as ``DVD player", ``lamp", or ``washing machine". This product description refers to the largest item purchased on a given visit to the store. For a separate sample (not overlapping with the main sample described above) covering purchases made between December 2008 and August 2009, both the basic product description and a product category assigned by the company is also available. I create a mapping between basic product descriptions and product categories in this sample and use it to assign a product category to each purchase in the 2005-2006 sample.[16] The product category is based on the company categorizing products into nine different departments, with each of the departments further sub-divided into classes. Purchases in the 2005-2006 sample fall into 124 product categories. Some of the product categories account for very small fractions of overall purchases. Within each of the nine departments, I therefore group some of the classes together and work with a total of 32 product categories.

The following unique feature of the lending process enables a straightforward study of the relation between what a customer purchases and default. Rather than having one revolving credit account at the company to which various purchases could be charged, the company issues a separate loan for each purchase. For example, suppose a customer buys a refrigerator and then comes back a few weeks later and buys an armchair. The company will make one loan for the refrigerator and another for the armchair and I am able to follow the repayment (or lack of repayment) of each of these two loans. Clothing and cell phone minutes are an exception to this principle since these are charged to a revolving credit account much like a U.S. credit card. This makes it difficult to compare losses on clothing and cell phone minutes to losses on other products and I therefore leave out clothing

[16] To create this mapping, I calculate (in the December 2008-August 2009 sample) the most common product category for each basic product description. For most basic product descriptions (73%) only one product category is used and for the remaining product descriptions the same product category is assigned by the company in the vast majority of the occurrences of a given basic product description. This implies that one with a high degree of accuracy can use this mapping to define product categories in the 2005-2006 sample by assigning the most common product category for that basic product description to all sales with a given basic product description.

and cell phone minutes from the majority of the analysis. The counts of individuals and purchases stated above exclude purchases of clothing and cell phone minutes.

The first column of Table 4 shows the distribution of sales (by peso value) across the nine departments, along with the fraction of sales constituted by clothes and cell phone minutes. Table 5 provides a more detailed breakdown across the 32 product categories, clothes and cell phone minutes. About 41% of sales are electronics, 31% are clothes, 9% are appliances, 8% are various types of furniture, 6% are kids' gear and toys or auto parts (of which kids items constitute more than half), 2% are kitchen equipment, with the remaining 3% constituted by watches, jewelry, eye glasses, and cell phone minutes. The range of products sold is thus very diverse, though one should keep in mind that consumer goods only represent a fraction of overall spending, with food and housing likely constituting a larger fraction for most households.

3.2 The mechanics of the loan process

The lending process for a particular purchase starts with the customer deciding which item(s) he or she would like to buy. A sales person then accompanies the client to the credit desk. For new clients a host of information is then collected, including the client's name, identity documentation, address, demographics, employer and income. If the client does not work, the spouse/partner's employment information is collected instead. The credit desk then verifies the clients identity, home address and work information by phone via a call center. This takes only a short period during which the client watches an informational video.

The credit desk then proposes a minimum down-payment. The rules for down-payments have changed over time, but the latest rules as of the end of the data set in 2009 are as follows.

% of customer's authorized credit	Type of client				
	A	B	N	C	D
From 0 to 100	0	20	10	30	30
101 to 150	10	20	20	30	40
151 to 200	20	30	30	40	50
201 to 300	20	40	40	50	60
301 to 400	30	50	50	60	70

Each cell states the required minimum down-payment as a percent of the cost of the item as a function of the cost of the item relative to the customer's authorized credit (credit limit) and the company's internal credit score for the client. A new customer's authorized credit (credit limit) is 25% of the customer's annual income. Subsequent limits are updated based on the client's payment history. A customer can borrow more than the limit but will then be required to pay a larger down-payment as laid out in the table above. The customer's internal credit score is calculated based on the customer's repayment efficiency to date. Repayment efficiency is calculated as the sum of actual payments divided by the sum of payments due since the customer first started borrowing at the company. Customers with repayment efficiencies above 75% are assigned a credit score of A, repayment efficiencies between 50% and 75% imply a credit score of B, repayment efficiencies between 25% and 50% imply a credit score of C, while repayment efficiencies below 25% imply a credit score of D. New customers are assigned a credit score of N, meaning that they have no repayment history.

The monthly payment on a loan is calculated as:

$$\text{Monthly payment}=\text{Loan amount}*(1+r)/\text{Loan term} \tag{6}$$

where r is the interest rate on the loan. The implied annual percentage rate on the loan is higher than r. For example, an interest rate of 24% on a 12-month loan leads to the same monthly payment using the above formula as an annual percentage rate of 41.6% with monthly compounding would. Interest rates are homogeneous across borrowers. Notably, they do not depend on the borrower's credit score, the down-payment, or the size of the purchase. The only variation in interest rates (at a given point in time) is that they are higher for cell phones than for other product categories, higher for 18-month loans than 12-month loans, and higher for cities considered high risk. The schedule of interest rates as of the end of the sample is:

City type:	Zone 1 (low risk)	Zone 2 (high risk)
Furniture/household item (12 month loan)	24%	30%
Furniture/household item (18 month loan)	36%	45%
Cell phone (12 month loan)	32%	38%
Cell phone (18 month loan)	44%	38%

Once the loan is granted, monthly bills are delivered by hand and explained in person. Additional visits are paid if the customer is overdue on his/her payments. If a customer misses payments, there are two possible outcomes: (1) The customer agrees to return the product to the store, or (2) the customer never pays what is owed and the firm declares the amount owed on the particular loan at this point a loss.

3.3 Differences in loan loss rates across product categories

When a customer does not make the full set of payments on a particular loan, the amount declared lost by the company is given by:

$$\text{Loss}=\text{Loan}*(1+r)-\text{Payments} \tag{7}$$

which can be decomposed into how much of the principal is not repaid and how much of the interest is not repaid:

$$\begin{aligned}\text{Loss}&=[\text{Loan}-\min(\text{Loan},\text{Payments})]+[\text{Loan}*r-(\text{Payments}-\min(\text{Loan},\text{Payments})]\\&=\text{Principal loss}+\text{Interest loss}.\end{aligned} \tag{8}$$

I define the loss rate as the loss divided by the size of the loan. One can approximate the company's realized return over the term of the loan as:

$$\begin{aligned}1+\text{realized return}&=\text{Payments}/\text{Loan}=(\text{Loan}*(1+r)-\text{Loss})/\text{Loan}\\&=(1+r)-[\text{Principal loss}/\text{Loan}]-[\text{Interest loss}/\text{Loan}].\end{aligned} \tag{9}$$

If all payments were due at the end of the payment term, this would be the exact realized return. Since in practice payments are due monthly, the realized return for the company (accounting for the fact that it can re-lend payments received before the end of the term) will be higher. On the other hand, if some of the payments received are made after the term of the loan these should be discounted back to the loan maturity date to calculate the exact realized return over the term of the loan. Ignoring this issue in the above approximation will tend to overstate the realized return.[17]

Table 4 amd 5 documents the main result from the Mexican data set: Dramatic differences in the company's loss rates and realized return on loans across products. Table 4, column (4) shows the lender's loss rate for each of the nine departments of products. The rates are calculated at the department level (as opposed to being an average across purchases in the category) in order to

[17] A related issue is that the return given in equation 4 is for the period of the term, as opposed to being on an annual basis. For the nine department that I focus my analysis on the loan term is either 12 or 18 months, but more than 99% of loans are 12 month loans. Dropping the 18 month loans from the analysis has very limited impact on any of the results.

account for any potential correlation between losses and purchase size. Four of the departments -- kitchen equipment, the two types of furniture (mattresses, dining sets and other furniture; living room and bedroom furniture), and appliances -- have substantially lower default rates than the others. For, these four departments, loss rates are between 11% and 12%. In contrast, loss rates for electronics, which constitute a large fraction of both sales and loans, have default rates above 20%. Loans given to finance jewelry purchases have default rates of almost 40% but jewelry constitute a small fraction of sales. Table 5 shows the loss rates for the 32 more detailed product categories. Consistent with the result for US data, product categories related to entertainment tend to have high default rates. I highlight these in pink.[18]

From the description of the loan process, it is clear that the differences in loss rates do not fully translate into corresponding differences in interest rates across products. The company does charge higher interest rates for cell phone loans, but otherwise charges the same interest rate for all products. Column (7) of Table 4 and 5 show the average interest rate charged for each department and for each of the 32 product categories. Rates are around 25% for each product category for the 2005-2006 sample analyzed here, with the exception of an average rate around 30% for cell phones.[19] Since a higher interest rate mechanically will lead to a higher interest loss rate for identical payments by the customer, column (5) and (6) decompose the loss rates into principal loss rates and interest loss rates. It is clear that even the principal loss rate (which is not mechanically affected by the loan interest rate) is higher for the entertainment-related categories than for most other categories.

Column (8) of Table 4 and 5 summarizes the impact of loss rates and interest rates on lender profits by showing the realized return earned by the lender for each type of product. The lender return is negative for loans given to finance jewelry purchases, due to the large loss rates for this category. For the other categories, the lender return is substantially positive since interest rates are far above loss rates, with patterns across departments and across product categories driven by the patterns in loss rates. It is important to emphasize that the lender return on loans calculated here do not account for the large expenses the company incurs as a result of employing thousands of staff to manage

[18] Phones (landline phones) are an exception, perhaps because they are purchased for businesses. Notice the low loss rate on office furniture.

[19] For the categories other than cell phones the small differences across categories are driven by slight differences in the timing of purchases (since interest rates change over time) and the location of purchases (across high and low risk cities).

the loan process. Accounting for differences in these expenses across product categories would likely increase differences in lender returns across product categories since additional costs from extra home visits are incurred when a customer starts missing payments on the loan. Limited available data on product markups across categories suggest that markups are not systematically related to loss rates.

Importantly, though loss rates are lower for more seasoned borrowers, Table 6 shows that the differences across product categories remain about as large in relative terms for seasoned as for new borrowers. For example, the default rate on electronics for loans to customers in their first month with the company is about 1.8 times the default rate on kitchen equipment, while the ratio of the default rates for these two categories is about 2.0 for loans to customers who have been with the company between 18 and 24 months at the time of purchase.

These findings indicate that the company could likely benefit from conditioning loan terms -- interest rates, down-payments, or credit limits -- much more on product type than was done over the sample. In fact, in early 2009, the firm increased the down-payment requirements for new clients from 10% to 20% for the following products: Cell phones, stereos, video games, iPods, computers, laptops, and jewelry.

To ensure that the differences in lender losses across product categories remain once these known predictors of default are controlled for, and to investigate how much additional predictive power is gained by considering product categories, I turn next to statistical models for predicting loss rates.

3.4. The predictive power of product mix for losses, controlling for standard default predictors

For a given loan, the loss rate (denote it by y) is either zero or positive. I am interested in how E(y|X) depends on a set of predictors X. One possible approach to modeling this relation would be to estimate a Tobit model. In that setup, E(y|X) would be non-linear in X (see Wooldridge (2002), equation (16.14)). I instead take a simpler approach and assume that E(y|X) is approximately linear in X over the relevant range of variation in X and proceed to estimate linear regression models by OLS. The reason for this simplification is that I later turn to estimating models with customer fixed effects. While one can estimate the regression coefficients (β's) in a Tobit model with fixed effects by transforming the model in a way that eliminates the fixed effects (see Honore (1992)), $dE(y|X)/dX$ (with fixed effects included in the set of X-variables) remains a function of the fixed effects and no unbiased estimator of the fixed effects exists. For comparability of results I therefore proceed to estimate linear regression models both for the cases without individual fixed effects and for the cases with individual fixed effects.

Table 7 predicts loss rates using time as customer dummies (to account for the strong negative relation between loss rates and time as a customer), transaction characteristics (loan amount, down payment/purchase price, interest rate, and loan term), measures of borrower credit risk (including the company's internal credit rating), demographics, and store fixed effects. For reference, Appendix Table 2 shows the summary statistics for the variables included in the regression.

The results in Table 7 show that both transactions characteristics, measures of credit risk, and demographics have explanatory power for predicting the loss rate on a given loan. Loan amount and loan interest rate both enter with positive signs. The effect of loan amount could reflect adverse selection (high-risk individuals self-select into larger loans) or moral hazard (a larger loan increases the likelihood of default either via strategic default or simply lack of affordability of the payments).[20] A positive coefficient on the interest rate could reflect the lender having information about likely losses and setting rates accordingly. Such interest variation is used by this particular company in the interest variation across cell phones versus other goods, across term of loan, across cities perceived by the company as low risk or high risk and across calendar time. Based on column (5), a one standard deviation (1309 peso) increase in loan amount increases the predicted loss rate by 1.1 percentage point, while a one standard deviation (3.9 percentage point) increase in the interest rate increases the predicted loss rate by 5.4 percentage points. Higher down payments are associated with a 0.9 percentage point decrease in the predicted loss rate for a one standard deviation (0.083) down payment/price increase. This effect could pick up differences across consumers in how long they have been planning for the purchase as well as a causal effect of ensuring lower monthly payments that are more affordable for the borrower. The term of the loan enters with a negative sign, possibly due to more affordable payments given the longer term.

[20] Adams, Einav and Levin (2009) argue, however, that one can include the excess of down-payments above a statistically predicted value as a proxy for an individual's risk type (since low risk borrowers use this to signal their type to affect the interest rate), in which case the effect of loan size measures only the moral hazard effect. This effect is fund to be positive. At the firm analyzed here, the interest rate on a loan does not depend on the down payment made thus preventing signaling so loan size effects could be due to either adverse selection or moral hazard.

As measures of credit risk I include dummies for the credit score groups used by the company. In addition to the credit score, I include the underlying repayment efficiencies for both the main account (I use main account to refer to loans for products in any of the nine departments, but remember that loans are made at the purchase level not the account level) and the clothing account. I include the number of purchases made to date as an additional risk control. I furthermore include variables that would matter for credit scores in the U.S. FICO score system and for which I have data: Credit limit, current amount of account balance, current amount of late balances, amount of moratory interest accumulated (and not paid) to date due to late or missing payments, and maximum credit level obtained in the past.[21] Customers with an A credit score are estimated to have loss rates 6 percentage points lower than customers with an N credit score. A larger number of purchases, large balances, late balances, or moratory interest are associated with higher predicted loss rates (likely due to indicating a larger financial strain imposed on the borrower relative to available resources), while large credit levels in the past are associated with lower predicted loss rates, possibly by indicating that the borrower has had the ability to repay large balances in the past. Of the demographics, age and years living at home address have the strongest relation to loss rates in economic terms (for a one standard deviation change). A one standard deviation (10.8 year) increase in age lowers the predicted loss rate by 2.3 percentage points, while a one standard deviation (11.3 year) increase in years living at home address lowers the predicted loss rate by 2.0 percentage points.[22]

Table 8, column (2)-(6) repeats regressions (1)-(5) from Table 3, but now adding dummies for the 32 product categories. The objective is first to determine whether the large differences in loss rates across product categories remain once transactions characteristics, credit risk measures, and demographics are controlled for, and second to determine how much incremental explanatory power the product category dummies add. For reference, Table 8, column 1, shows a regression of loss rates on only the product category dummies themselves. In order to focus on differences in loss rates across product categories, I pick the category with the lowest default rate, sewing

[21] These variables are as of the end of the month prior to the month of the loan analyzed, or as of the date of the first loan for customers getting their first loan, to make sure they are observable at the time of the loan.

[22] For a given loan I use the demographics as of the end of the prior month, or as of the date of the first loan for customers getting their first loan. The only exception is that the household size variables are only available as of December 2008. Results are largely unaffected by excluding the household size measures. The company restricts credit for minors, so I include a dummy for being a minor (age<21 for men, age<18 for women) in addition to age in the regressions.

machines, as a reference category (omitted dummy) and show the dummies on the other product categories which then measure how much higher the average loss rate is for a given product category relative to the average loss rate for loans for sewing machines. The table indicates significance levels for the product dummies by using a smaller and italic font for coefficients that are not significant at the 5% level.[23]

Moving from column (2) to column (6) in Table 8, I add still more regressors as indicated in the top part of the table. Controlling for time as customer fixed effects, transactions characteristics, measures of borrower credit risk, demographics, and store fixed effects has very little effect on the relative differences in loss rates across product categories. For example, audio not for cars have an average loss rate that is 12.2 percentage points higher than the average loss rate for sewing machines when no controls are included, and the difference is still as high as 9.6 percentage points when including the controls listed above. The large differences in loss rates across product categories are thus to a large extent robust to controlling for standard predictors of default.

In terms of explanatory power, the product category dummies on their own generate an R^2 of 0.021. Comparing column (2)-(6) of Table 8 to column (1)-(5) of Table 7 allows for an evaluation of the incremental R^2 from adding the product dummies. In each case, the R^2 in Table 8 is between 0.01 and 0.02 higher with the product dummies. While this is small in absolute terms, it is economically meaningful given the fact that R^2-values in regressions that predict loss rates for consumer credit tend to be very small both for the company analyzed here and in prior work (e.g., Gross and Souleles (2002)).

3.5 Product effects versus individual effects

To help assess the economic forces underlying differential lender losses across product categories, I estimate models of loss rates that include customer fixed effects. This allows us to assess whether differential loss rates across product categories are driven mainly by which types of individuals buy particular products or whether loss rates then to be high in certain products regardless of who buys them.

[23] The relative magnitudes differ a bit from those of the loss rates presented in Table 5 because those loss rates were at the product category level, while the product dummy coefficients in the regressions estimate the average loss rates across loans in a given category (relative to the benchmark) and thus implicitly weights each loan equally regardless of its size.

Consider a decomposition in which the loss rate on a loan made to individual i for buying product p has a product-specific component and an individual-specific component as well as a component driven by observables

$$Loss\ rate_{i,p} = f_p + f_i + x_{i,p}'\beta. \tag{10}$$

The average loss rate across I individuals borrowing for purchasing product p is then

$$Average\ Loss\ Rate_p = \frac{1}{I}\sum_{i=1}^{I} Loss\ rate_{i,p} = f_p + \left(\frac{1}{I}\sum_{i=1}^{I} f_i\right)_p + \frac{1}{I}\sum_{i=1}^{I} x_{i,p}'\beta \tag{11}$$

In this setting, if one estimates a loan level regression for loss rates, including product dummies and observables, the regression coefficient on the product dummy for product category p will estimate $f_p + \left(\frac{1}{I}\sum_{i=1}^{I} f_i\right)_p$. It will thus capture both the product-specific effect for product category p and the average individual-specific component for individuals taking out loans to purchase products in product category p. If one instead estimate the same regression, but now include both product dummies, individual dummies (individual fixed effects), and observables, then the regression coefficient on the product dummy for product category p will estimate only the product-specific effect f_p.

In order to be able to identify the product-specific effect f_p by including individual fixed effects in the regression, it is necessary that a lot of individual make purchases across several categories of products. Of 499,906 customers represented in the regressions, 179,311 purchased goods in both one or more of the four departments in Table 4 with lowest loss rates and in one or more of the five departments in Table 4 with the highest loss rates. Furthermore, focusing on the 32 more detailed product categories, the difference between the lowest and highest (of 32) default categories purchased by a given customer is 5.5% on average across customers. This suggests that there should be sufficiently many individuals with purchases across both high and low loss categories to separately identify the impact of product-specific effects and individual-specific effects.

Column (7) of Table 8 adds individual fixed effects to the product-level loss rate regression. The impact on the coefficients for the product category dummies is dramatic. The majority of them are now economically small and 12 of them are not significant.

Figure 2 illustrates the impact of including individual fixed effects on the regression coefficients for the product category dummies. The figure sorts the 32 product categories based on their average loss rate from column (1) of Table 8. These average loss rates are illustrated by the upward sloping line in the figure, with each point labeled with the number of the product category used in Table 5. The flatter line in the figure illustrates the coefficients on the product category dummies from column (7) of Table 8, i.e. the ``true'' product-specific effects once the impact of which customers tend to buy particular products is taken out. Most of the product-specific effects are economically small. The vertical difference between the two lines shows the average individual-specific component for individuals taking out loans to purchase products in the category (plus the small effect of the observables). The vertical differences are large implying that there are large risk differences across products in the risk of the customer pool they attract.

The importance of person fixed effects is consistent with my interpretation of the US results as suggesting that particular products tend to attract particular people, though the Mexican data do not allow further analysis of the nature of the cross-person heterogeneity.

4. Conclusion

In this paper, I study what lenders can learn from account-level data. I focus on analysis of spending patterns across consumption categories. By shedding light on why account-level data are useful, the paper contributes to understanding the big data revolution in credit scoring and the economic fundamentals of credit demand and credit risk. My central finding is that spending patterns across goods and services are informative for consumer credit outcomes. In particular, spending on entertainment-related goods and services is related to a higher probability of paying positive finance/interest/late charges (in US data) or of not repaying consumer credit in full (in Mexican data). I conjecture that higher entertainment-related spending is an indicator of impatience and show in the US data that the same spending categories that predict paying positive finance/interest/late charges also predict smoking and lower education, outcomes typically associated with impatience.

References

Adams, William, Liran Einav and Jonathan Levin, 2009, ``Liquidity Constraints and Imperfect Information in Subprime Lending'', *American Economic Review*, 99, 1, 49-84.

Banerjee, Abhijit and Sendhil Mullainathan, 2010, ``The Shape of Temptation: Implications for the Economic Lives of the Poor", NBER working paper 15973.

Bellenger, Danny N., Dan H. Robertson, and Elizabeth C. Hirschman, 1978, ``Impulse Buying Varies by Product", *Journal of Advertising Research*, 18, 6, 15-18.

Berg, T., V. Burg, A. Gombovic and M. Puri, 2020, ``On the Rise of FinTechs: Credit Scoring Using Digital Footprints", *Review of Financial Studies*, 33, 7, 2845-2897.

Calvet, L. E., J. Y. Campbell, F. Gomes and P. Sodini, 2021, "The Cross-Section of Household Preferences", CEPR Discussion Paper DP16105.

Charles, Kerwin Kofi, Erik Hurst and Nikolai Roussanov, 2002, ``Conspicuous Consumption and Race", *Quarterly Journal of Economics*, 124, 2, 425-467.

Dhar, Ravi and Klaus WertenBroch, ``Consumer Choice between Hedonic and Utilitarian Goods", *Journal of Marketing Research*, 37, 1, 60-71.

Fair Isaac Corporation, 2005, ``Understanding Your FICO Score".

Gathergood, J., 2012, "Self-control, financial literacy and consumer over-indebtedness", *Journal of Economic Psychology*, 33, 590–602.

Gross, David B. and Nicholas S. Souleles, 2002, ``An Empirical Analysis of Personal Bankruptcy and Delinquency", *Review of Financial Studies*, 15, 1, 319-347.

Honore, Bo. E, 1992, ``Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects", *Econometrica*, 60, 533-567.

Iyer, R., A. I. Khwaja, E. P. Luttmer and K. Shue, 2016, "Screening Peers Softly: Inferring the Quality of Small Borrowers", *Management Science*, 62, 6, 1554-1577.

Keys, Benjamin J., Tanmoy K. Mukherjee, Amit Seru and Vikrant Vig, 2010, ``Did Securitization Lead to Lax Screening", *Quarterly Journal of Economics*, 125, 1, 307-362.

Khan, Uzma, Ravi Dhar, and Klaus Wertenbroch, 2004, ``A Behavioral Decision Theoretic Perspective on Hedonic and Utilitarian Choice", INSEAD working paper, 2004.

Laibson, David, 1997, ``Golden Eggs and Hyperbolic Discounting", *Quarterly Journal of Economics*, May, 443-477.

Puri, Radhika, 1996, ``Measuring and Modifying Consumer Impulsiveness: A Cost-Benefit Accessibility Framework", *Journal of Consumer* Psychology, 5, 2, 87-113.

Rook, Dennis W. and Robert J. Fisher, 1995, ``Normative Influences on Impulsive Buying Behavior", *Journal of Consumer Research*, 22, 3, 305-313.

Vissing-Jorgensen, Annette, 2002, ``Limited Asset Market Participation and the Elasticity of Intertemporal Substitution", *Journal of Political Economy*, 110, 4, 825-853.

Wooldridge, Jeffrey M., 2002, *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

World Bank, 2019, Credit Scoring Approaches Guideline.

Table 1. US data: Relating consumer credit finance/interest/late charges to spending mix. Univariate approach (with controls).

Each row in the table presents the results of linear regressions of a dummy for paying positive finance/interest/late charges over the past 12 months on consumer credit on a dummy for positive spending in a given category and a set of controls (equation (1)). Consumer credit includes major credit cards, store credit cards, gasoline credit cards, store installment credit accounts and credit from financial institutions. It excludes mortgages, home equity loans, vehicle loans, and business related loans. Controls include the log and log squared of total real expenditure, the log and log squared of after-tax real income, the respondent's age and age squared, family size, a dummy for the respondent being male, a dummy for the households residing in a rural area, and time dummies (quarterly). Each regression is based on data for 63,691 unique households over the period 1988Q1-2013Q1. Results are shown for the 50 spending categories with the largest t-statistic for γ_c (in absolute value). Spending categories related to other borrowing or banking are highlighted in light blue. Spending categories related to entertainment are highlighted in pink.

Spending category	t-statistic for γ_c	Coefficient (γ_c)	Fraction of households with positive spending on category
CHECK ACCTS / OTH BANK SERV CHGS	25.53	0.094	0.438
VEHICLE INSURANCE	13.93	0.070	0.819
AUTOMOBILE FINANCE CHARGES	13.83	0.055	0.318
MORTGAGE INTEREST OWND	12.07	0.049	0.518
RNTL VIDEO CASS/TAPES/DISCS/FILMS	11.69	0.047	0.553
TRUCK FINANCE CHARGES	11.30	0.051	0.219
VIDEO CASSETTES/TAPES/DISCS	9.02	0.035	0.402
COOLANT/ADDITIVES/BRK/TRNS FLD	8.90	0.037	0.273
PHYSICIANS SERVICES	8.73	0.034	0.613
WOMENS HOSIERY	8.71	0.035	0.430
PRESCRIPTION DRUGS	8.59	0.036	0.699
VET SERVICES	8.39	0.034	0.290
INTEREST, LMP SUM HM EQ LN, OWND	8.27	0.068	0.056
MAGAZINE/NEWSPAPER SINGLE COPY	8.19	0.057	0.328
PET-PURCHASE/SUPPLIES/MEDICINE	8.17	0.031	0.430
TOYS GAMES ARTS CRAFTS TRICYCLES	8.17	0.031	0.493
VIDEO GAME HARDWARE/SOFTWARE	8.11	0.040	0.186
SCHOOL MEALS	7.87	0.039	0.223
LOTTERIES AND PARIMUTUEL LOSSES	7.66	0.044	0.276
RECORDS,CDS,AUDIO TAPES	7.66	0.034	0.473
RCRD/TAPE/CD/VIDEO MAIL ORD CLUB	7.38	0.048	0.121
FILM	7.37	0.031	0.443
CIGARETTES	7.20	0.028	0.329
WOMENS ACCESSORIES	7.01	0.029	0.305
MAGAZINES, NON-SUBSCRIPTION	6.87	0.037	0.374
WOMENS UNDERGARMENTS	6.79	0.026	0.405
NEWSPAPERS, NON-SUBSCRIPTION	6.77	0.036	0.416
BOOKS NOT THRU BOOK CLUBS	6.71	0.025	0.507
BATHROOM LINENS	6.64	0.028	0.255
WOMENS SHIRTS, TOPS,BLOUSES	6.54	0.025	0.580

RENTERS INSURANCE	6.16	0.047	0.060
TELEPHONE SERVICE NOT SPEC	6.15	0.146	0.960
GIRLS ACCESSORIES	6.10	0.040	0.086
JEWELRY	5.99	0.024	0.344
OTHER GASH GIFTS	5.99	0.032	0.368
GIRLS SWIMSUITS/WARM-UP/SKI SUITS	5.92	0.037	0.097
MENS HOSIERY	5.87	0.025	0.269
INT PAID ON OTH VEH	5.81	0.066	0.026
OTH HOUSEHOLD DECORATIVE ITEMS	5.80	0.025	0.350
WOMENS PANTS	5.66	0.024	0.521
WOMENS FOOTWEAR	5.60	0.021	0.550
BOOKS THRU BOOK CLUBS	5.60	0.032	0.114
CREDIT CARD MEMBERSHIPS	5.58	0.039	0.085
ELECTRIC PERSONAL CARE APPL.	5.44	0.027	0.153
COMPTER SFTWR/CMPTR ACC N-BUS USE	5.40	0.029	0.146
WOMENS DRESSES	5.39	0.021	0.365
SCHOOL BK/SUPL/EQUIP FOR ELEM/HS	5.29	0.030	0.138
	…		
ELECTRICITY OWNV	-5.79	-0.072	0.022
CAP IMPROVE LABOR/MAT OWND	-6.04	-0.031	0.158
SAFE DEPOSIT BOX RENTAL	-9.96	-0.060	0.109

Table 2. US data: Relating consumer credit finance/interest/late charges to spending mix. Multivariate approach: LASSO

The table presents the result of estimation of the multivariate model in equation (3), estimated using LASSO (with λ chosen by the extended bayesian information criterion). The table presents the spending mix variables chosen by LASSO, sorted based on the post-estimation OLS t-statistics. The t-statistic for γ_c and the coefficients for γ_c are based on the post-estimation OLS regression. These are biased and are not to be used for inference but as indicators of the relative importance of variables. The controls chosen by LASSO are shown at the bottom of the table.

Spending category	t-statistic for γ_c	γ_c
CHECK ACCTS / OTH BANK SERV CHGS	21.23	0.079
MORTGAGE INTEREST OWND	17.00	0.081
AUTOMOBILE FINANCE CHARGES	11.36	0.045
ELECTRICITY RNTR	10.07	0.053
TRUCK FINANCE CHARGES	9.41	0.042
VEHICLE INSURANCE	9.25	0.046
INTEREST, LMP SUM HM EQ LN, OWND	7.84	0.064
CIGARETTES	5.66	0.022
CAR LEASE PAYMENTS	5.11	0.049
MAGAZINE/NEWSPAPER SINGLE COPY	4.67	0.032
RNTL VIDEO CASS/TAPES/DISCS/FILMS	4.65	0.019
RENTERS INSURANCE	4.27	0.034
PHYSICIANS SERVICES	4.06	0.016
COIN-OP HSHLD LNDRY, DRY CLN	3.97	0.023
RCRD/TAPE/CD/VIDEO MAIL ORD CLUB	3.95	0.026
INT PAID ON OTH VEH	3.82	0.043
WOMENS HOSIERY	3.56	0.015
SCHOOL MEALS	3.52	0.017
LOTTERIES AND PARIMUTUEL LOSSES	3.43	0.019
VET SERVICES	3.35	0.015
NEWSPAPERS, NON-SUBSCRIPTION	3.35	0.018
COOLANT/ADDITIVES/BRK/TRNS FLD	3.22	0.014
BATHROOM LINENS	2.60	0.011
VIDEO GAME HARDWARE/SOFTWARE	2.34	0.012
MAGAZINES, NON-SUBSCRIPTION	2.27	0.013
GIRLS ACCESSORIES	2.05	0.015
RECORDS,CDS,AUDIO TAPES	1.97	0.009
VIDEO CASSETTES/TAPES/DISCS	1.90	0.008
GIRLS SWIMSUITS/WARM-UP/SKI SUITS	1.74	0.012
WOMENS ACCESSORIES	1.48	0.006
TOYS GAMES ARTS CRAFTS TRICYCLES	1.23	0.005
PET-PURCHASE/SUPPLIES/MEDICINE	0.69	0.003
ELECTRICITY OWNV	-4.39	-0.053
DOMESTIC SERVICE	-5.56	-0.034
CASH CONTRIBUTIONS TO EDUCATIONAL INSTITUTIONS	-5.76	-0.054
MEDICARE PAYMENTS	-6.00	-0.030

CAP IMPROVE LABOR/MAT OWND	-6.68	-0.034
SAFE DEPOSIT BOX RENTAL	-9.75	-0.057
ln(Real after-tax income)	3.23	0.006
D(Male)	-5.37	-0.020
D(Rural)	-5.79	-0.038

Time dummies included without penalty, coefficients omitted for brevity

Table 3. US data: Does spending mix capture time preferences?

The table figures relates categories' ability to predict positive finance/interest/late change on consumer credit (γ_c) to categories' ability to predict smoking (γ_c^{smoke}) or years of education ($\gamma_c^{education}$).

Column (1) has two fewer spending categories since it omits the two categories (cigarettes and other tobacco spending) using to define smoking.

	Dependent variable:			
	t-statistic for γ_c^{smoke}	γ_c^{smoke}	t-statistic for $\gamma_c^{education}$	$\gamma_c^{education}$
	(1)	(2)	(3)	(4)
t-statistic for γ_c	0.663***		-0.705***	
	(7.46)		(-5.33)	
γ_c		0.624***		-2.231***
		(7.17)		(-5.64)
Fraction of households with positive spending on category	-7.545***	-0.0370***	14.44***	0.172***
	(-5.16)	(-4.00)	(6.62)	(4.09)
Constant	-1.337***	-0.00874***	3.548***	0.0867***
	(-3.60)	(-3.64)	(6.40)	(7.94)
N (spending categories)	446	448	448	448
R^2	0.124	0.110	0.108	0.079

Table 4. Mexican data: Loss rates by product category

	Pct. of sales	Excluding products with no default information (clothes, cell phone minutes)						
		Pct. of sales	Pct. of loans	Loss rate	Principal loss rate	Interest loss rate	Average interest rate charged	Lender return =(7)-(4)
Product category	(1)	(2)	(3)	**(4)**	(5)	(6)	(7)	**(8)**
Kitchen equipment, various hh. items	2.4%	3.5%	3.5%	**11.5%**	7.8%	3.7%	24.9%	**13.4%**
Electronics	40.6%	60.0%	60.2%	**21.3%**	13.8%	7.5%	27.6%	**6.3%**
Mattresses, dining sets, other furniture	4.9%	7.2%	7.2%	**11.3%**	7.4%	3.8%	24.9%	**13.6%**
Living room and bedroom furniture	3.4%	5.1%	5.0%	**11.1%**	7.0%	4.1%	25.7%	**14.6%**
Kids gear and toys, auto parts, bikes	5.5%	8.2%	8.3%	**16.5%**	11.1%	5.4%	24.9%	**8.4%**
Appliances	9.2%	13.5%	13.4%	**11.8%**	7.4%	4.4%	25.5%	**13.7%**
Watches	0.6%	0.9%	0.9%	**17.0%**	11.6%	5.5%	25.0%	**8.0%**
Jewelry	0.7%	1.1%	1.1%	**39.2%**	27.7%	11.5%	25.2%	**-14.0%**
Eye glasses etc.	0.3%	0.5%	0.5%	**15.4%**	10.3%	5.2%	25.0%	**9.6%**
Cell phone minutes	1.8%							
Clothes	30.5%							
All above categories	100%	100%	100%	**18.2%**	11.8%	6.4%	26.6%	**8.4%**

Table 5. Mexican data: Loss rates by detailed product category

Product category	Pct. of sales (1)	Excluding products with no default information (clothes, cell phone minutes): Pct. of sales (2)	Pct. of loans (3)	Loss rate (4)	Principal loss rate (5)	Interest loss rate (6)	Average interest rate charged (7)	Lender return =(7)-(4) (8)
Kitchen equipment, various household items								
1. Kitchen electronics	1.3%	1.9%	1.9%	**11.1%**	7.5%	3.6%	25.0%	13.9%
2. Cook and tableware	0.4%	0.6%	0.6%	**11.8%**	8.0%	3.8%	24.9%	13.1%
3. Personal care	0.3%	0.5%	0.5%	**13.3%**	9.0%	4.3%	24.8%	11.5%
4. Luggage	0.3%	0.4%	0.4%	**12.2%**	8.4%	3.9%	24.6%	12.4%
Electronics								
5. Audio, for cars	3.0%	4.4%	4.5%	**20.5%**	14.0%	6.5%	25.0%	4.4%
6. Audio, not for cars	5.6%	8.2%	8.2%	**16.2%**	10.7%	5.6%	25.8%	9.6%
7. TVs	5.0%	7.4%	7.4%	**18.7%**	12.5%	6.2%	25.4%	6.7%
8. DVD, video	2.3%	3.4%	3.5%	**15.8%**	10.6%	5.1%	25.2%	9.4%
9. Entertainment electronics	2.9%	4.2%	4.2%	**18.5%**	12.7%	5.8%	25.0%	6.4%
10. Phones (not cell)	0.4%	0.6%	0.6%	**9.7%**	6.5%	3.3%	25.1%	15.4%
11. Cell phones	20.8%	30.8%	31.0%	**24.9%**	15.8%	9.1%	29.7%	4.8%
12. Microwave ovens	0.5%	0.8%	0.8%	**13.5%**	9.0%	4.5%	25.1%	11.7%
Mattresses, dining sets, other furniture								
13. Mattresses	2.2%	3.2%	3.2%	**12.6%**	8.4%	4.2%	24.9%	12.4%
14. Dining sets, chairs	1.1%	1.7%	1.6%	**11.3%**	7.4%	3.9%	24.8%	13.5%
15. Office furniture	0.2%	0.3%	0.3%	**7.7%**	5.1%	2.6%	24.9%	17.2%
16. Wardrobes, cupboards	1.0%	1.4%	1.4%	**8.4%**	5.5%	2.9%	24.8%	16.4%
Living room and bedroom furniture								
17. Living room furniture	2.6%	3.9%	3.8%	**11.4%**	7.1%	4.2%	25.8%	14.5%
18. Bedroom furniture	0.5%	0.7%	0.7%	**12.1%**	7.6%	4.5%	25.3%	13.2%
19. Sewing machines	0.3%	0.5%	0.5%	**7.3%**	4.9%	2.4%	25.1%	17.8%
Kids gear and toys, auto parts, bikes								
20. Baby items (e.g. stroller)	0.9%	1.4%	1.4%	**17.6%**	11.8%	5.8%	24.9%	7.3%
21. Toys	0.8%	1.2%	1.2%	**17.2%**	11.9%	5.4%	24.9%	7.7%
22. Tires, car batteries	2.1%	3.0%	3.1%	**16.0%**	10.6%	5.4%	24.8%	8.8%
23. Kids bikes	1.7%	2.5%	2.5%	**16.4%**	11.2%	5.2%	24.9%	8.5%
Appliances								
24. Fans, AC units	0.9%	1.4%	1.4%	**13.9%**	9.0%	4.9%	25.0%	11.1%
25. Water heaters, other heaters	0.4%	0.6%	0.6%	**11.3%**	7.4%	3.9%	25.3%	14.0%
26. Stoves, ovens	1.7%	2.5%	2.5%	**11.4%**	7.3%	4.1%	25.2%	13.8%
27. Fridges, water coolers	3.0%	4.4%	4.3%	**12.4%**	7.6%	4.8%	25.7%	13.3%

28. Washer/dryer/dishwasher	3.1%	4.6%	4.6%	**10.9%**	6.7%	4.1%	25.5%	14.7%
29. Other (from above categories)	0.7%	1.1%	1.1%	**11.2%**	7.3%	3.9%	24.9%	13.7%
30. Watches	0.6%	0.9%	0.9%	**17.0%**	11.6%	5.5%	25.0%	8.0%
31. Jewelry	0.7%	1.1%	1.1%	**39.2%**	27.7%	11.5%	25.2%	-14.0%
32. Glasses etc.	0.3%	0.5%	0.5%	**15.4%**	10.3%	5.2%	25.0%	9.6%
33. Cell phone minutes	1.8%	0.0%						
34. Clothes	30.5%	0.0%						
All above categories	100%	100%	100%	**18.2%**	11.8%	6.4%	26.6%	8.4%

Table 6. Mexican data: Loss rate by product category and time as customer

Product category	Loss rate, by months as customer at time of current purchase				
	<1	1 to 6	6 to 12	12 to 18	18 to 24
Kitchen equipment, various household items	14.3%	12.9%	6.9%	6.7%	8.3%
Electronics	25.4%	22.4%	14.3%	14.2%	16.3%
Mattresses, dining sets, other furniture	12.6%	12.5%	8.7%	8.4%	10.1%
Living room and bedroom furniture	11.8%	12.6%	9.1%	8.8%	11.0%
Kids gear and toys, auto parts, bikes	19.0%	19.1%	10.7%	11.4%	13.9%
Appliances	13.2%	13.3%	8.7%	8.7%	10.1%
Watches	21.6%	18.4%	9.4%	10.4%	12.1%
Jewelry	51.5%	34.6%	20.1%	24.0%	28.7%
Eye glasses etc.	18.5%	16.7%	9.4%	10.2%	14.6%
All above categories	21.5%	19.6%	12.3%	12.2%	14.4%

Table 7. Mexican data: Predicting loss rates using information known at time of purchase

Significance indicated with *** (1%), ** (5%) and * (10%).

	Dependent variable: Loss rate=Amount not repaid/Loan amount				
	(1)	(2)	(3)	(4)	(5)
Fixed effects (month dummies) for time as customer	Yes	Yes	Yes	Yes	Yes
Transaction characteristics					
Loan amount (1000s of pesos)		0.0100***	0.0100***	0.0093***	0.0087***
Downpayment/Purchase price		-0.0525***	-0.1339***	-0.1240***	-0.1132***
Interest rate		1.0649***	1.0260***	0.9892***	1.3823***
Term of loan (months)		-0.0133***	-0.0131***	-0.0123***	-0.0205***
Measures of borrower credit risk					
Credit score (omitted: New customer, no score)					
A (best credit)			-0.0621***	-0.0618***	-0.0597***
B			0.0292***	0.0252***	0.0250***
C			0.0674***	0.0625***	0.0642***
D			0.1546***	0.1554***	0.1543***
Repayment efficiency, main account			-0.0012***	-0.0011***	-0.0011***
Repayment efficiency, clothing account			-0.0011***	-0.0011***	-0.0010***
Credit limit (omitted: limit=4200 pesos)					
Limit=8400 pesos			-0.0055***	-0.0086***	0.0003
Limit=12600 pesos			-0.0082***	-0.0029**	0.0163***
Number of purchases made to date			0.0094***	0.0080***	0.0080***
Account balance, main account (1000s of pesos)			0.0188***	0.0186***	0.0178***
Account balance, clothing account (1000s of pesos)			0.0646***	0.0664***	0.0641***
Late balance, main account (1000s of pesos)			0.1091***	0.1048***	0.1037***
Late balance, clothing account (1000s of pesos)			0.0905***	0.0870***	0.0869***
Moratory interest accumulated, main account (1000s of pesos)			0.6919***	0.6505***	0.6374***
Moratory interest accumulated, clothing account (1000s of pesos)			0.7327***	0.7259***	0.7196***
Maximum credit level in the past, main account (1000s of pesos)			-0.0112***	-0.0109***	-0.0108***
Maximum credit level in the past, clothing account (1000s of pesos)			-0.0093***	-0.0080***	-0.0079***

Demographics					
Age				-0.0021***	-0.0021***
Minor (age<21 for men, age<18 for women)				-0.0011	0.0011
Male				0.0255***	0.0214***
Marital status (omitted: married)					
Divorced				0.0672***	0.0650***
Single				0.0140***	0.0114***
Couple, not married				0.0374***	0.0359***
Widow				0.0395***	0.0376***
Income category (omitted: income<4200 pesos)					
>=4200, <8400 pesos				-0.0118***	-0.0124***
>=8400, <12600 pesos				-0.0099***	-0.0117***
>=12600, <16800 pesos				-0.0132***	-0.0115***
>=16800 pesos				-0.0104***	-0.0097***
Highest education (omitted: no schooling)					
<=Elementary school				0.0096***	0.0013
<=Junior high				0.0154***	-0.0016
<=Technical college				-0.0050*	-0.0275***
<=High school				0.0110***	-0.0086***
<=University				-0.0217***	-0.0414***
Living situation (omitted: home owner)					
Renter				0.0540***	0.0534***
Lives with family				0.0072***	0.0040***
Guest				0.0040	0.0059
Years living at home address				-0.0016***	-0.0018***
Number of people living in customer's house				-0.0076***	-0.0072***
Number of people who live in customer's house and work				0.0138***	0.0141***
Number of people who are economically dependent on the client				-0.0001	0.0003
Store fixed effects	No	No	No	No	Yes
N	1,364,864	1,364,864	1,364,864	1,364,864	1,364,864
R2	0.015	0.027	0.068	0.084	0.097

Table 8. Mexican data: Predicting loss rates using information known at time of purchase, including product categories

	Dependent variable: Loss rate=Amount not repaid/Loan amount						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Controls:							
Time as customer fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes
Transactions characteristics	No	No	Yes	Yes	Yes	Yes	Yes
Measures of borrower credit risk	No	No	No	Yes	Yes	Yes	Yes
Demographics	No	No	No	No	Yes	Yes	Yes
Store fixed effects	No	No	No	No	No	Yes	No
Individual fixed effects	No	No	No	No	No	No	Yes
Product category (omitted=sewing machines)							
Kitchen equipment, various household items							
Kitchen electronics	0.030	0.029	0.054	0.052	0.049	0.048	0.008
Cook and tableware	0.025	0.026	0.052	0.05	0.048	0.049	0.015
Personal care	0.047	0.047	0.074	0.065	0.056	0.053	0.009
Luggage	0.038	0.041	0.067	0.06	0.057	0.058	0.009
Electronics							
Audio, for cars	0.122	0.122	0.121	0.110	0.100	0.096	0.028
Audio, not for cars	0.074	0.074	0.080	0.071	0.064	0.061	*0.006*
TVs	0.102	0.099	0.090	0.083	0.074	0.072	*0.000*
DVD, video	0.078	0.075	0.087	0.078	0.067	0.063	*0.003*
Entertainment electronics	0.089	0.096	0.092	0.080	0.071	0.068	*0.007*
Phones (not cell)	0.018	0.020	0.043	0.037	0.038	0.031	*0.005*
Cell phones	0.167	0.163	0.151	0.138	0.128	0.129	0.043
Microwave ovens	0.053	0.054	0.067	0.062	0.056	0.051	*0.009*
Mattresses, dining sets, other furniture							
Mattresses	0.044	0.044	0.045	0.043	0.042	0.039	0.021
Dining sets, chairs	0.030	0.033	0.027	0.023	0.020	0.019	*-0.001*
Office furniture	*-0.003*	*-0.004*	*0.011*	*0.009*	*0.012*	*0.009*	0.013
Wardrobes, cupboards	*0.009*	*0.011*	0.016	0.015	*0.010*	*0.009*	0.010
Living room and bedroom furniture							
Living room furniture	0.026	0.027	*0.004*	*0.000*	*-0.002*	*-0.003*	*-0.006*
Bedroom furniture	0.032	0.033	*-0.013*	*-0.014*	-0.020	-0.019	-0.019
Sewing machines	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*	*0.000*
Kids gear and toys, auto parts, bikes							
Baby items (e.g. stroller)	0.090	0.087	0.106	0.096	0.086	0.084	0.023
Toys	0.095	0.095	0.114	0.099	0.090	0.088	0.019
Tires, car batteries	0.083	0.086	0.101	0.094	0.097	0.093	0.033
Kids bikes	0.084	0.083	0.092	0.080	0.075	0.077	0.011

Appliances							
Fans, AC units	0.047	0.046	0.061	0.056	0.051	0.050	*0.000*
Water heaters, other heaters	0.036	0.041	0.049	0.046	0.049	0.038	0.015
Stoves, ovens	0.040	0.041	0.035	0.034	0.030	0.031	*-0.001*
Fridges, water coolers	0.044	0.042	*0.004*	*0.002*	*-0.006*	*-0.006*	-0.022
Washer/dryer/dishwasher	0.029	0.028	*0.007*	*0.005*	*0.000*	*-0.001*	-0.015
Other (from above categories)	0.017	0.021	0.037	0.033	0.033	0.032	*0.005*
Watches	0.075	0.076	0.096	0.084	0.079	0.080	0.025
Jewelry	0.186	0.184	0.202	0.184	0.175	0.175	0.031
Eye glasses etc.	0.079	0.078	0.091	0.084	0.084	0.081	0.036
N=1,364,864							
R2	0.021	0.035	0.039	0.077	0.092	0.103	0.819

Figure 1. US data: Does spending mix capture time preferences?

The figures relate categories' ability to predict positive finance/interest/late change on consumer credit (γ_c) to categories' ability to predict smoking (γ_c^{smoke}) or years of education ($\gamma_c^{education}$). The figures present bin-scatter plots with 10 bins.

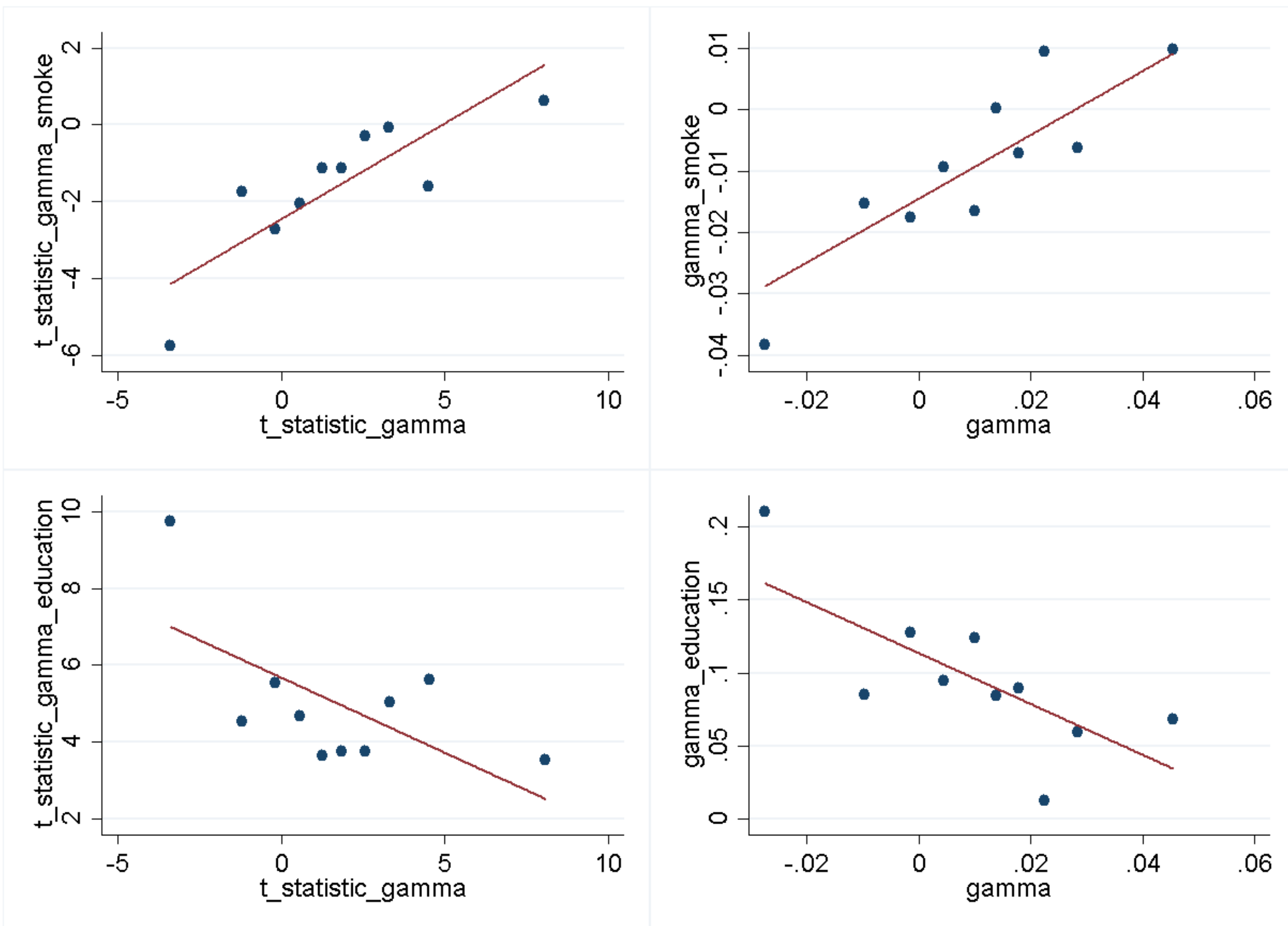
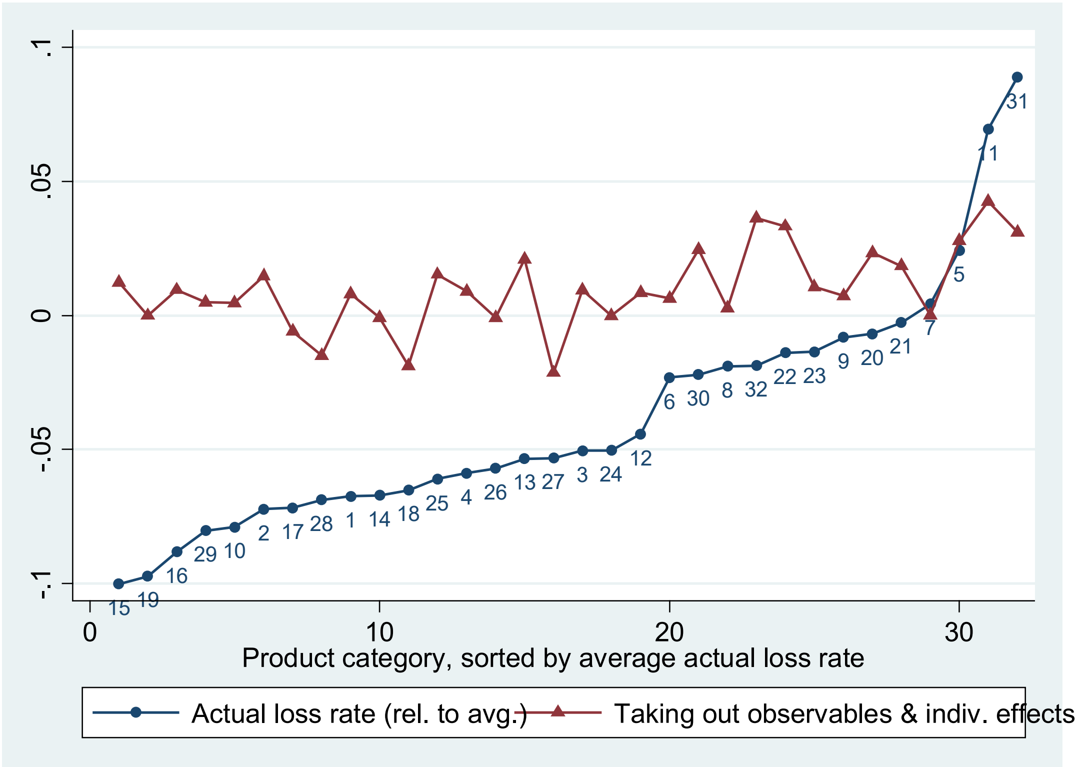


Figure 2. Mexican data: Product or individual effect? Average loss rates with and without individual fixed effects



Appendix Table 1. The 448 spending categories used from the US Consumer Expenditure Survey

FOOD OR BOARD WHILE AT SCHOOL
FOOD AND BEV FOR CATERED AFFAIRS
FOOD OUT OF TOWN TRIPS
FOOD PREPARED BY CU ON TRIPS
ALC. BEV. PURCHASED ON TRIPS
RENT OF DWELLING
LODGING AWAY FROM HOME
HOUSING FOR SOMEONE AT SCHOOL
GROUND RENT OWND
FIRE/EXTENDED COVERAGE OWND
HOMEOWNERS INSURANCE OWND
HOMEOWNERS INSURANCE OWNV
PROPERTY TAXES OWND
PROPERTY TAXES OWNV
MORTGAGE INTEREST OWND
MORTGAGE INTEREST OWNV
INTEREST, LMP SUM HM EQ LN, OWND
CAPITAL IMPROVE MATERIALS OWND
CAP IMPROVE LABOR/MAT OWND
CAP IMPROVE LABOR/MAT OWNV
PARKING OWND
PAINTING/PAPERING LABOR/MAT OWND
PLMB/WTR HEAT LABOR/MAT OWND
HEAT/AC/ELEC LABOR/MAT OWND
ROOFING/GUTTERS LABOR/MAT OWND
OTH REP/MAINT LABOR/MAT OWND
REPL DISHWASH/DISP/HOOD OWND
HRD SURFACE FLOOR LABOR/MAT OWND
W/W CARPET INST REPL OWND
REPAIR-DISPL/DWSHR/RANG HD OWND
REP/MAINT LABOR/MAT RNTR
OTH REP/MAINT LABOR/MAT OWND
OTH REP/MAINT LABOR/MAT OWNV
PROP MANAGEMENT OWND
PROP MANAGEMENT OWNV
PAINT/WALLPAPER AND SUPP RNTR
PAINT/WALLPAPER AND SUPP OWND
EQUIP FOR PAINT/WPAPER RNTR
EQUIP FOR PAINT/WPAPER OWND
MAT FOR PANEL/SIDING, ETC OWND
MAT/EQUIP FOR ROOF/GUTTER OWND
MAT FOR PATIO,MASONRY,ETC OWND
PLUMBING SUPP/EQUIP OWND
ELEC SUPP,HEAT/COOL EQUIP OWND
FUEL OIL OWND
GAS, BOTTLED OR TANK OWND
GAS,BOTTLED OR TANK - OWNV & RVS
WOOD/KEROSENE/OTHER FUELS OWND
ELECTRICITY RNTR
ELECTRICITY OWND

ELECTRICITY OWNV
UTILITY--NATURAL GAS RNTR
UTILITY--NATURAL GAS OWND
UTILITY--NATURAL GAS OWNV
TELEPHONE SERVICE NOT SPEC
RESIDENTIAL TELEPHONES/PAY PHONES
CELLULAR PHONE SERVICE
PHONE CARDS
WATER AND SEWERAGE MAINT RNTR
WATER AND SEWERAGE MAINT OWND
WATER AND SEWERAGE MAINT OWNV
CABLE/SATELLITE/COM ANTENNA SERV
SATELLITE RADIO SERVICE
TRASH/GARBAGE COLLECT RNTR
TRASH/GARBAGE COLLECT OWND
TRASH/GARBAGE COLLECT OWNV
SEPTIC TANK CLEANING OWND
BATHROOM LINENS
BEDROOM LINENS
KITCHEN AND DINING ROOM LINENS
CURTAINS AND DRAPES
SLIPCOVERS/DECORATIVE PILLOWS
SEWING MATERIALS
OTHER LINENS
MATTRESS AND SPRINGS
OTHER BEDROOM FURNITURE
SOFAS
LIVING ROOM CHAIRS
LIVING ROOM TABLES
KITCHEN/DINING ROOM FURNITURE
INFANTS FURNITURE
OUTDOOR FURNITURE
WALL UNITS, CABINETS, OCCAS FURN
PURCH/INST REFRIG/FREEZER RNTR
PURCH/INST REFRIG/FREEZER OWND
PURCH/INST CLOTHES WASHER RNTR
PURCH/INST CLOTHES WASHER OWND
PURCH/INST CLOTHES DRYER RNTR
PURCH/INST CLOTHES DRYER OWND
STOVES, OVENS OWND
MICROWAVE OVENS RNTR
MICROWAVE OVENS OWND
PURCH/INST WINDOW A/C OWND
COLOR TV - CONSOLE
COLOR TV - PORTABLE/TABLE MOD
TELEVISIONS
VCRS/VIDEO DISC PLAYERS
VIDEO CASSETTES/TAPES/DISCS
VIDEO GAME HARDWARE/SOFTWARE
VIDEO GAME SOFTWARE
STREAMING/DOWNLOADING VIDEO
RADIOS
TAPE RECORDERS AND PLAYERS
DIGITAL AUDIO PLAYERS

COMPONENTS/COMPONENT SYSTEMS
ACCESSORIES AND OTH SOUND EQUIP
ACCESSORIES AND OTHER SOUND EQUIP
RECORDS,CDS,AUDIO TAPES
RCRD/TAPE/CD/VIDEO MAIL ORD CLUB
RECORDS,CDS,AUDIO TAPES
STREAMING/DOWNLOADING AUDIO
FLOOR COVERINGS (NON-PERM.)
FLOOR COVERINGS (NON-PERM.)
WINDOW COVERINGS
INFANTS EQUIPMENT
BARBEQUE GRILLS AND OUTDOOR EQUIP
CLOCKS
LAMPS AND LIGHTING FIXTURES
OTH HOUSEHOLD DECORATIVE ITEMS
TELEPHONES AND ACCESSORIES
CLOCKS AND OTHER HH DECOR ITEMS
PLASTIC DINNERWARE
CHINA AND OTHER DINNERWARE
FLATWARE
GLASSWARE
OTHER SERVING PIECES
NONELECTRIC COOKWARE
LAWN AND GARDEN EQUIPMENT
POWER TOOLS
ELECTRIC FLOOR CLEANING EQUIP
SEWING MACHINES
SMALL ELECTRIC KITCHEN APPLIANCES
PORTABLE HEATING/COOLING EQUIP
CONSTRUCTION MAT OWND
FLOOR REPAIR/REPL MATERIALS OWND
LANDSCAPING MATERIALS OWND
OFFICE FURNITURE HOME USE
HAND TOOLS
INDOOR PLANTS, FRESH FLOWERS
CLOSET AND STORAGE ITEMS
MAT FOR TERMTE/PST CNTRL MAINTCE
BABYSITTING
BABYSIT/CHILD CARE OWN HOME
BABYSIT/CHILD CARE OTHER HOME
DOMESTIC SERVICE
GARDENING/LAWN CARE SERVICE
WATER SOFTENING SERVICE
MOVING, STORAGE,FREIGHT
HSHLD LNDRY,DRYCLN NOT COIN-OP
COIN-OP HSHLD LNDRY, DRY CLN
REPAIR OF TV/RADIO/SOUND EQUIP
REPAIR OF HOUSEHOLD APPLIANCES
REUPHOLSTERY OF FURNITURE
RENTAL/REPAIR-TOOLS,LAWN/GARDEN
MISC. HOME SERVICES
RENTAL OF HOUSEHOLD EQUIPMENT
MNGMT/SPEC SER/SECURITY OWND
SERV FOR TERMT/PST CNTRL

HOME SECURITY SYS. SERV. FEE
RENTERS INSURANCE
MENS SUITS
MENS SPORTCOATS/TAILORED JACKETS
MENS COATS AND JACKETS
MENS UNDERWEAR
MENS HOSIERY
MENS NIGHTWEAR/LOUNGEWEAR
MENS ACCESSORIES
MENS SWEATERS AND VESTS
MENS SWIMSUITS/WARM-UP/SKI SUITS
MENS SHIRTS
MENS PANTS
MENS SHORTS/SHORTS SETS
MENS PANTS AND SHORTS
MENS UNIFORMS
MENS COSTUMES
BOYS COATS AND JACKETS
BOYS SWEATERS
BOYS SHIRTS
BOYS UNDERWEAR
BOYS NIGHTWEAR
BOYS HOSIERY
BOYS ACCESSORIES
BOYS SUITS, SPORTCOATS,VESTS
BOYS PANTS
BOYS SHORTS, SHORTS SETS
BOYS PANTS AND SHORTS
BOYS UNIFORMS/ACTIVE SPORTSWE
BOYS COSTUMES
BOYS UNIFORMS
BOYS SWIMSUITS/WARM-UP/SKI SUITS
WOMENS COATS AND JACKETS
WOMENS DRESSES
WOMENS SPORTCOATS, TAIL. JKTS
WOMENS VESTS AND SWEATERS
WOMENS SHIRTS, TOPS,BLOUSES
WOMENS SKIRTS
WOMENS PANTS
WOMENS SHORTS,SHORTS SETS
WOMENS PANTS AND SHORTS
WOMENS SWIMSUITS/WARM-UP/SKI SUIT
WOMENS SLEEPWEAR
WOMENS UNDERGARMENTS
WOMENS HOSIERY
WOMENS SUITS
WOMENS ACCESSORIES
WOMENS UNIFORMS
WOMENS COSTUMES
GIRLS COATS AND JACKETS
GIRLS DRESSES, SUITS
GIRLS SHIRTS/BLOUSES/SWEATERS
GIRLS SKIRTS AND PANTS
GIRLS SHORTS, SHORTS SETS

GIRLS SKIRTS, PANTS, AND SHORTS
GIRLS SWIMSUITS/WARM-UP/SKI SUITS
GIRLS UNDERWEAR AND SLEEPWEAR
GIRLS HOSIERY
GIRLS ACCESSORIES
GIRLS UNIFORMS
GIRLS COSTUMES
MENS FOOTWEAR
BOYS FOOTWEAR
GIRLS FOOTWEAR
WOMENS FOOTWEAR
INFANT COAT/JACKET/SNOWSUIT
INFANT COAT/JACKET/SNOWSUIT 9B
INFANT DRESSES/OUTERWEAR
INFANT DRESSES/OUTERWEAR 9B
INFANT UNDERGARMENTS
INFANT UNDERGARMENTS 9B
INFANT NIGHTWEAR/LOUNGEWEAR
INFANT NIGHTWEAR/LOUNGEWEAR 9B
INFANTS ACCESSORIES
MATERIAL FOR MAKING CLOTHES
SEWING NOTIONS, PATTERNS
WATCHES
JEWELRY
LUGGAGE
SHOE REPAIR, OTH SHOE SERVICE
COIN-OP APPAREL LDRY/DRY CLNG
ALTER/REPAIR OF APPAREL, ACCESS
CLOTHING RENTAL
WATCH AND JEWELRY REPAIR
APPAREL LNDRY/DRY CLNG N/COIN-OP
NEW CARS
TRADE-IN ALLOWANCE/NEW CARS
NEW TRUCKS
TRADE-IN ALLOW/NEW TRUCKS
CAR LEASE PAYMENTS
TRUCK LEASE PAYMENTS
USED CARS
TRADE-IN ALLOWANCE/USED CARS
USED TRUCKS
TRADE-IN ALLOWANCE/USED TRUCKS
GASOLINE
DIESEL FUEL
GASOLINE ON OUT OF TOWN TRIPS
MOTOROIL
MOTOR OIL ON OUT OF TOWN TRIPS
COOLANT/ADDITIVES/BRK/TRNS FLD
TIRES PURCHASED/REPLACED/INSTALL
PARTS/EQUIP/ACCESSORIES
VEHICLE PRODUCTS & SERVICES
PARTS/EQUIP/ACCESSORIES
BODY WORK AND PAINTING
CLUTCH, TRANSMISSION REPAIR
DRIVE SHAFT AND REAR-END REPAIR

BRAKE WORK
BRAKE WORK
REPAIR TO STEERING OR FRONT END
REPAIR TO ENGINE COOLING SYSTEM
MOTOR TUNE-UP
LUBE, OIL CHANGE AND OIL FILTERS
FRNT END ALIGN, WHEEL BAL/ROTAT
SHOCK ABSORBER REPLACEMENT
BRAKE ADJUSTMENT
TIRE REPAIR AND OTH REPAIR WORK
VEHICLE AIR CONDITION REPAIR
EXHAUST SYSTEM REPAIR
ELECTRICAL SYSTEM REPAIR
MOTOR REPAIR/REPLACEMENT
VEHICLE ACCESSORIES INCL. LABOR
VEHICLE AUDIO EQ. INCL. LABOR
AUTO REPAIR SERVICE POLICY
VEHICLE INSURANCE
AUTOMOBILE FINANCE CHARGES
TRUCK FINANCE CHARGES
MOTORCYCLE & PLANE FINANCE CHG
VEHICLE REGISTRATION STATE/LOCAL
State vehicle registration
Local vehicle registration
DRIVERS LICENSE
VEHICLE INSPECTION
AUTO RENTAL
AUTO RENTAL, OUT-OF-TOWN TRIPS
TRUCK RENTAL
TRUCK RENTAL, OUT-OF-TOWN TRIP
PARKING FEES
PRKNG FEE IN HME CITY EXCL RSDNC
PARKING FEES, OUT-OF-TOWN TRIP
TOLLS OR ELECTRONIC TOLL PASSES
TOLLS ON OUT-OF-TOWN TRIPS
TOWING CHARGES
DOCKING/LANDING FEES
AIRLINE FARES
INTERCITY BUS FARES
INTRACITY MASS TRANSIT FARES
LOCAL TRANS. OUT OF TOWN TRIPS
TAXI FARES ON TRIPS
TAXI FARES AND LIMOUSINE SERVICE
INTERCITY TRAIN FARES
SHIP FARES
PRESCRIPTION DRUGS
EYEGLASSES AND CONTACT LENSES
MEDICAL EQUIP. FOR GENERAL USE
SUPPORTIVE/CONVAL MED. EQUIP.
HEARING AIDS
PHYSICIANS SERVICES
DENTAL SERVICES
EYECARE SERVICES
SERV BY PRCTIONER OTH THAN PHYS

LAB TESTS, X-RAYS
SERV BY PROS OTH THAN PHYSICIANS
HOSPITAL ROOM
HOSPITAL ROOMS & SERVICES
HOSPITAL SERVICE OTH THAN ROOM
OTHER MEDICAL CARE SERVICE
RENTAL OF MEDICAL/SURGICAL EQUIP
RENTAL OF SUPORTIVE/CONVAL EQUIP
COMMERCIAL HEALTH INSURANCE
TRD FEE FOR SRV HLTH P (NO BCBS)
TRD FEE FOR SRV HLTH P (BCBS)
PREF PROVIDER HLTH PLN (NO BCBS)
PREF PROVIDER HLTH PLN (BCBS)
BLUECROSS/BLUE SHIELD
HEALTH MAINTENANCE PLANS
HLTH MAINT. ORG (NO BCBS)
HLTH MAINT. ORG (BCBS)
LONG TERM CARE INSURANCE
MEDICARE PAYMENTS
COML MEDICAR SUPLMNT/OTH HLTH INS
COML MEDICARE SUPPLEMNT (NO BCBS)
COML MEDICARE SUPPLEMENT (BCBS)
OTHER HEALTH INSURANCE (NO BCBS)
MEDICARE PRES. DRUG PREMIUMS
NEWSPAPERS
NEWSPAPER SUBSCRIPTIONS
NEWSPAPERS, NON-SUBSCRIPTION
MAGAZINES
MAGAZINE SUBSCRIPTIONS
MAGAZINES, NON-SUBSCRIPTION
BOOKS THRU BOOK CLUBS
BOOKS NOT THRU BOOK CLUBS
MAGAZINE/NEWSPAPER SUBSCRIPTION
MAGAZINE/NEWSPAPER SINGLE COPY
GENERAL SPORT/EXCERCISE EQUIP
BICYCLES
CAMPING EQUIPMENT
HUNTING, FISHING EQUIPMENT
WINTER SPORT EQUIPMENT
WATER SPORT/OTHER SPORT EQUIP
WATER SPORT EQUIPMENT
OTHER SPORT EQUIPMENT
TOYS GAMES ARTS CRAFTS TRICYCLES
PLAYGROUND EQUIPMENT
MUSIC INSTRUMENTS/ACCESSORIES
FILM
PHOTOGRAPHIC EQUIPMENT
PET-PURCHASE/SUPPLIES/MEDICINE
REC EXPNS OUTSIDE HOME CITY
CLUB MEMBERSHIP DUES AND FEES
SOCIAL/RECRE/CIVIC CLUB MEMBRSHP
CREDIT CARD MEMBERSHIPS
AUTOMOBILE SERVICE CLUBS
SHOPPING CLUB MEMB FEES

FEES FOR PARTICIPANT SPORTS
PARTIC. SPORTS OUT-OF-TOWN TRI
MOVIE, THEATER, OPERA, BALLET
MOVIE, OTH ADM. OUT-OF-TOWN
ADMISSION TO SPORTING EVENTS
ADM TO SPRTS EVENTS OUT-OF-TOW
FEES FOR RECREATIONAL LESSONS
PHOTOGRAPHER FEES
FILM PROCESSING
PET SERVICES
VET SERVICES
OTH ENT SERV, OUT-OF-TOWN TRIP
RENT/REP MUSIC INSTRUMENTS
RENT/REPAIR OF MISC SPORTS EQU
RNTL VIDEO CASS/TAPES/DISCS/FILMS
LOTTERIES AND PARIMUTUEL LOSSES
ONLINE ENTERTAINMENT AND GAMES
CIGARETTES
OTHER TOBACCO PRODUCTS
WIGS AND HAIRPIECES
ELECTRIC PERSONAL CARE APPL.
PERS. CARE SERV FOR FEMALES
PERS. CARE SERV FOR MALES
PERS. CARE SERV.
SCHOOL BK/SUPL/EQUIP FOR COLLEGE
SCHOOL BK/SUPL/EQUIP FOR ELEM/HS
ENCYL. OTH SETS OF REFRNCE BKS
SCH BKS/SUPP-DAY CARE,NURS,OTH
COLLEGE TUITION
ELEM./H.S. TUITION
DAY CARE/NURS/PRSCH EXP INCL TUIT
OTHER SCHOOL TUITION
OTH SCH EXPENSES INCLUD RENTALS
LEGAL FEES
FUNERAL EXPENSE
SAFE DEPOSIT BOX RENTAL
CHECK ACCTS / OTH BANK SERV CHGS
CEMETERY LOTS,VAULTS,MAINT FEES
ACCOUNTING FEES
COMPTER/COMPTER HRDWAR N-BUS USE
COMPTER SFTWR/CMPTR ACC N-BUS USE
REPAIR-CMPTR,CMPTR SYS N-B
COMPUTER INFORMATION SERVICES
INTERNET SERVICES AWAY FROM HOME
PORTABLE MEMORY
COMPUTER ACCESSORIES
TELEPHONE ANSWERING DEVICES
CALCULATORS
TYPWRITS/OTH OFF MACH NON-BUS USE
SMOKE ALARM PUR/RENT OWND
OTH HH APPL RNTR
OTH HH APPL OWND
REGULAR GROC SHOPPING INCL GOODS
FOOD/NONALC BEV AT GROC STORES

FD/NONALC BEV AT CONVEN STORE
AVG FOOD/NONALC BEV EXPENSES
BEER/WINE FOR HOME USE
OTHER ALCOHOL FOR HOME USE
BEER/WINE/OTH ALC FOR HOME USE
DINING OUT AT REST., ETC EXCL ALC
ALCOHOL AT RESTAURANTS ETC
SCHOOL MEALS
MAINT/REP/UTIL OTH PROP
Child support expenditures
Rent received as pay
CSH GFT/NON-CU, CNTRB/ORG
Support for college students
Cash contributions to charities, other organizatio
Cash contributions to churches or religious organi
Cash contributions to educational institutions
Cash contributions to political organizations
Other cash gifts
INT PAID ON OTH VEH
INTEREST, HM EQ LN (CRDT), OWND

Appendix Table 2. Summary statistics for the Mexican data set

	N (number of non-missing obs.)	10th percen-tile	50th percen-tile	90th percen-tile	Mean	Std. Dev.
Loss rate	1,364,864	0	0	0.937	0.170	0.385
Time as a customer (months)	1,364,864	0	1	13	4.093	5.383
Transaction characteristics						
Loan amount (1000s of pesos)	1,364,864	0.349	1.258	3.199	1.585	1.309
Down payment/Purchase price	1,364,864	0.100	0.104	0.205	0.132	0.083
Interest rate	1,364,864	0.240	0.300	0.360	0.305	0.039
Term of loan (months)	1,364,864	12	12	12	12.055	0.553
Measures of borrower credit risk						
Credit score						
N (new customer)	1,316,623	1	1	1	0.944	0.229
A (best credit)	1,316,623	0	0	0	0.049	0.217
B	1,316,623	0	0	0	0.005	0.071
C	1,316,623	0	0	0	0.001	0.026
D	1,316,623	0	0	0	0.000	0.022
Repayment efficiency, main account	690,776	77	100	106	96.518	23.001
Repayment efficiency, clothing account	651,259	67	100	100	93.203	24.728
Credit limit						
Limit=4200 pesos	1,287,424	0	0	1	0.342	0.474
Limit=8400 pesos	1,287,424	0	0	1	0.416	0.493
Limit=12600 pesos	1,287,424	0	0	1	0.243	0.429
Number of purchases made to date	1,364,864	1	2	6	2.911	2.767
Account balance, main account (1000s of pesos)	1,364,864	0	0	3.788	1.214	1.949
Account balance, clothing account (1000s of pesos)	1,364,864	0	0	1.182	0.355	0.679
Late balance, main account (1000s of pesos)	1,364,864	0	0	0	0.033	0.175
Late balance, clothing account (1000s of pesos)	1,364,864	0	0	0	0.025	0.128
Moratory interest accumulated, main account (1000s of pesos)	1,364,864	0	0	0	0.003	0.022
Moratory interest accumulated, clothing account (1000s of pesos)	1,364,864	0	0	0	0.002	0.015
Maximum credit level in the past, main account (1000s of pesos)	1,364,864	0	0.830	5.602	1.960	2.639
Maximum credit level in the past, clothing account (1000s of pesos)	1,364,864	0	0.216	1.868	0.643	0.924

Demographics						
Age	1,363,516	21.010	30.300	48.055	32.679	10.766
Minor (age<21 for men, age<18 for women)	1,085,905	0	0	0	0.025	0.155
Male	1,086,073	0	0	1	0.483	0.500
Marital status						
Married	1,086,069	0	1	1	0.540	0.498
Divorced	1,086,069	0	0	0	0.022	0.148
Single	1,086,069	0	0	1	0.311	0.463
Couple, not married	1,086,069	0	0	1	0.111	0.314
Widow	1,086,069	0	0	0	0.015	0.122
Income category						
<4200 pesos	1,364,548	0	1	1	0.556	0.497
>=4200, <8400 pesos	1,364,548	0	0	1	0.145	0.353
>=8400, <12600 pesos	1,364,548	0	0	1	0.104	0.305
>=12600, <16800 pesos	1,364,548	0	0	0	0.080	0.272
>=16800 pesos	1,364,548	0	0	1	0.114	0.318
Highest education						
No schooling	1,361,473	0	0	0	0.018	0.132
<=Elementary school	1,361,473	0	0	1	0.208	0.406
<=Junior high	1,361,473	0	0	1	0.351	0.477
<=Technical college	1,361,473	0	0	0	0.084	0.277
<=High school	1,361,473	0	0	1	0.199	0.399
<=University	1,361,473	0	0	1	0.141	0.348
Living situation						
Home owner	1,050,752	0	1	1	0.803	0.398
Renter	1,050,752	0	0	0	0.066	0.248
Lives with family	1,050,752	0	0	1	0.130	0.337
Guest	1,050,752	0	0	0	0.001	0.028
Years living at home address	1,363,355	2	10	28	13.324	11.258
Number of people living in customer's house	1,364,864	2	4	7	4.315	1.757
Number of people who live in customer's house and work	1,364,864	1	2	4	2.154	1.291
Number of people who are economically dependent on the client	1,364,864	0	2	4	1.782	1.847