Learning in the Credit Card Market

Sumit Agarwal, John C. Driscoll, Xavier Gabaix, and David Laibson*

February 8, 2008

Abstract

Agents with more experience make better choices. We measure learning dynamics using a panel with four million monthly credit card statements. We study add-on fees, specifically cash advance, late payment, and overlimit fees. New credit card accounts generate fee payments of $15 per month. Through negative feedback – i.e. paying a fee – consumers learn to avoid triggering future fees. Paying a fee last month reduces the likelihood of paying a fee in the current month by about 40%. Controlling for account fixed effects, monthly fee payments fall by 75% during the first three years of account life. We find that learning is not monotonic. Knowledge effectively depreciates about 10% per *month*, implying that learning displays a strong recency effect.

JEL classification: D1, D4, D8, G2.

Keywords: household finance, consumer finance, credit cards, feedback, learning, learning-by-doing, recency.

*Agarwal: Federal Reserve Bank of Chicago, sagarwal@frbchi.org. Driscoll: Federal Reserve Board, john.c.driscoll@frb.gov. Gabaix: New York University and NBER, xgabaix@stern.nyu.edu. Laibson: Harvard University and NBER, dlaibson@harvard.edu. Gabaix and Laibson acknowledge support from the National Science Foundation (Human and Social Dynamics program). Laibson acknowledges financial support from the National Institute on Aging (R01-AG-1665). The views expressed in this paper are those of the authors and do not represent the policies or positions of the Board of Governors of the Federal Reserve System or the Federal Reserve Bank of Chicago. Ian Dew-Becker, Keith Ericson and Tom Mason provided outstanding research assistance. The authors are grateful to Devin Pope for his discussion of the paper at the AEA, as well as Murray Carbonneau, Stefano Dellavigna, Joanne Maselli, Nicola Persico, Matthew Rabin, and seminar participants at the AEA, Berkeley, EUI and the NBER (Law and Economics) for their suggestions. This paper previously circulated under the title "Stimulus and Response: The Path from Naiveté to Sophistication in the Credit Card Market."

1 Introduction

Economists believe that learning through experience underpins optimization and generates technological progress. Large literatures measure learning dynamics in the lab,[1] and in the field.[2]

However, because of data limitations, relatively few papers measure learning in the field with micro-level (household) panel data. Among such household studies, most show that households learn to optimize over time. For example, Miravete (2003) and Agarwal, Chomsisengphet, Liu and Souleles (2007) respectively show that consumers switch telephone calling plans and credit card contracts to minimize monthly bill payments.

Moreover, a few papers are able to identify the specific information flows that elicit learning. For instance, Fishman and Pope (2006) study video stores, and find that renters are more likely to return their videos on time if they have recently been fined for returning them late. Ho and Chong (2003) use grocery store scanner data to estimate a model in which consumers learn about product attributes. They find that the model has greater predictive power, with fewer parameters, than forecasting models used by retailers.[3]

In this current paper, we study the process by which individual households learn to avoid add-on fees in the credit card market.[4] We analyze a panel dataset that contains three years of credit card statements, representing 120,000 consumers and 4,000,000 credit card statements. We focus our analysis on credit card fees — late payment, over limit, and cash advance fees — since some observers argue that new customers do not optimally minimize such fees.[5] We want to know whether credit card holders pay fewer fees with experience.

[1] For example, Van Huyck, Battalio and Beil (1990, 1991), McAllister (1991), Crawford (1995), Roth and Erev (1995), VanHuyck, Battalio and Rankin (2001), Anderson (2000), Camerer (2003), and Wixted (2004a, 2004b).

[2] For example, see Zimmerman (1982), Argote, Beckman and Epple (1990), Gruber (1992), Bahk and Gort (1993), Marimon and Sunder (1994), McAfee and McMillan (1996), Nye (1996), Sargent (1999), Benkard (2000), Thompson (2001), Thornton and Thompson (2001), Evans, Honkaphoja and Marimon (2001), Evans and Honkaphoja (2001), Barrios and Strobl (2004), List (2003), Harrison and List (2004).

[3] Lemieux and MacLeod (2000) study the effect of an increase in unemployment benefits in Canada. They find that the propensity to collect unemployment benefits increased with a first-time exposure to this new system via an unemployment spell. Barber, Odean and Strahlevitz (2004) find evidence that individual investors tend to repurchase stocks that they previously sold for a gain.

[4] There is a large literature on the magnitude of interest payments and fees in the credit card market: Ausubel (1991), Calem and Mester (1995), Massoud, Saunders and Scholnick (2006), Ausubel (1999), Kerr and Dunn (2002), Shui and Ausubel (2004), DellaVigna and Malmendier (2004), Kerr (2004), Calem, Gordy and Mester (2005).

[5] For example, Frontline reports that "The new billions in revenue reflect an age-old habit of human be-

We find that fee payments are very large immediately after the opening of an account. We find that new accounts generate direct monthly fee payments that average \$15 *per month*.[6] However, these payments fall by 75 percent during the first four years of account life. To formally study these dynamics, we estimate a learning model with the Method of Simulated Moments. The data reveal that learning is driven by feedback. Making a late payment – and consequently paying a fee – reduces the probability of another late payment in the subsequent month by 44 percent.

These learning effects may be driven by many different channels. Consumers learn about fees when they are forced to pay them. Alternatively, consumers may pay more attention to their credit card account when they have recently paid fees. Through these many channels, card holders learn to sharply cut their fee payments over time.

We find that the learning dynamics are not monotonic. Card holders act as if their knowledge depreciates – i.e., their learning patterns exhibit a recency effect.[7] A late payment charge from the previous month is more influential than an identical charge that was paid a year ago. The monthly hazard rate of a fee payment *increases* as previous fee payments recede further into the past (holding all else equal). We estimate that this knowledge effectively at a rate of between 10 and 20 percent per month. At first glance, this finding seems counter-intuitive. But there are actually several examples of papers that have found such forgetting effects. For instance, Benkard (2000) finds evidence for both learning and forgetting – that is, depreciation of productivity over time – in the manufacturing of aircraft, as do Argote, Beckman and Epple (1990), in shipbuilding.

Our findings imply that learning is very powerful, but that knowledge depreciation partially offsets learning. Nevertheless, the net effect of learning is clear. Learning generates a

havior: Most people never anticipate they will pay late, so they do not shop around for better late fees." (http://www.pbs.org/wgbh/pages/frontline/shows/credit/more/rise.html) There is also a nascent academic literature that studies how perfectly rational firms interact in equilibrium with imperfectly rational consumers. See Shui and Ausubel (2004), DellaVigna and Malmendier (2004, 2006), Miao (2005), Mullainathan and Shleifer (2005), Oster and Morton (2005), Gabaix and Laibson (2006), Heidhues and Koszegi (2006), Jin and Leslie (2006), Koszegi and Rabin (2006), Spiegler (2006), and Ellison (forthcoming) for an overview.

[6] Moreover, this understates the impact of fees, since some behavior — e.g. a pair of late payments — not only triggers direct fees but also triggers an interest rate increase, which is *not* captured in our \$15 calculation. Suppose that a consumer is carrying \$2,000 of debt. Changing the consumer's interest rate from 10% to 20% is equivalent to charging the consumer an extra \$200. Late payments also may prompt a report to the credit bureau, adversely affecting the card holder's credit accessability and creditworthness. The average consumer has 4.8 cards and 2.7 actively used cards.

[7] See Lehrer (1988), Piccione and Rubinstein (1997) and Aumann, Hart and Perry (1997) for some theoretical models of forgetfulness.

substantial net reduction in fee payments.

We organize our paper as follows, Section 2 summarizes our data and presents our basic evidence for learning and backsliding. Section 3 presents a model for those patterns. This model is estimated with the Method of Simulated Moments in section 4. Section 5 discusses alternative explanations for our findings. In Section 6, we draw some conclusions.

2 Two Patterns in Fee Payment

In this section, we describe the dataset and present two sets of reduced-form analyses. This analysis provides a summary of key moments. Formal estimation of a structural model follows in section 4.

2.1 Data

We use a proprietary panel dataset from a large U.S. bank that issues credit cards nationally. The dataset contains a representative random sample of about 128,000 credit card accounts followed monthly over a 36 month period (from January 2002 through December 2004). The bulk of the data consists of the main billing information listed on each account's monthly statement, including total payment, spending, credit limit, balance, debt, purchase and cash advance annual percentage rate (APR), and fees paid. At a quarterly frequency, we observe each customer's credit bureau rating (FICO score) and a proprietary (internal) credit 'behavior' score. We have credit bureau data about the number of other credit cards held by the account holder, total credit card balances, and mortgage balances. We have data on the age, gender and income of the account holder, collected at the time of account opening. Further details on the data, including summary statistics and variable definitions, are available in the data appendix.

We focus on three important types of fees, described below: late fees, over limit fees, and cash advance fees.[8]

[8]Other types of fees include annual, balance transfer, foreign transactions, and pay by phone. All of these fees are relatively less important to both the bank and the borrower. Fewer issuers (the most notable exception being American Express) continue to charge annual fees, largely as a result of increased competition for new borrowers (Agarwal et al., 2005). The cards in our data do not have annual fees. A balance transfer fee of 2-3% of the amount transferred is assessed on borrowers who shift debt from one card to another. Since few consumers repeatedly transfer balances, borrower response to this fee will not allow us to study learning about fee payment- though see Agarwal et al. (2006) for a discussion of other borrower uses of

1. **Late Fee**: A *direct* late fee of $30 or $35 is assessed if the borrower makes a payment beyond the due date on the credit card statement. If the borrower is late by more than 60 days once, or by more than 30 days twice within a year, the bank may also impose *indirect* late fees by raising the APR to over 24 percent.[9] Such indirect fees are referred to as 'penalty pricing.' The bank may also choose to report late payments to credit bureaus, adversely affecting consumers' FICO scores. Our data analysis covers only direct late fees (and not penalty pricing).

2. **Over Limit Fee**: A direct over limit fee, also of $30 or $35, is assessed the first time the borrower exceeds his or her credit limit in a given month. Penalty pricing also results from over limit transactions. Our data analysis covers only direct over limit fees (and not penalty pricing).

3. **Cash Advance Fee**: A direct cash advance fee of 3 percent of the amount advanced or $5 (whichever is greater) is levied for each cash advance on the credit card. Unlike the first two types of fees, a cash advance fee can be assessed many times per month. Cash advances do not invoke penalty pricing. However, the APR on cash advances is typically greater than that on purchases, and is usually 16 percent or more. Our data analysis covers only direct cash advance fees (and not subsequent interest charges).

2.2 Reduced form analyses

2.2.1 Fee payment by account tenure

Figure 1 reports the frequency of each fee type as a function of account tenure. The regression — like all those that follow — controls for time effects, account fixed effects, and time-varying attributes of the borrower (e.g. variables that capture card utilization each pay

balance transfer cards. The foreign transaction fees and pay by phone fees together comprise less than three percent of the total fees collected by banks.

[9] If the borrower does not make a late payment during the six months after the last late payment, the APR will revert to its normal (though not its promotional) level.

cycle). The data plotted in Figure 1 is generated by estimating,

$$
\begin{aligned}
(1) \qquad f^{j}_{i,t} &= \alpha + \phi_i + \psi_{time} + Spline(Tenure_{i,t}) \\
&+ \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} \\
&+ \gamma_1 Util_{i,t-1} + \epsilon_{i,t}.
\end{aligned}
$$

The dependent variable $f^{j}_{i,t}$ is a dummy variable that takes the value 1 if a fee of type j is paid by account i at tenure t. Fee categories, j, include late payment fees — $f^{Late}_{i,t}$ — over limit fees — $f^{Over}_{i,t}$ — and cash advance fees — $f^{Advance}_{i,t}$. Parameter α is a constant; ϕ_i is an account fixed effect; ψ_{time} is a time fixed-effect; $Spline(Tenure_{i,t})$ is a spline[10] that takes account tenure (time since account was opened) as its argument; $Purchase_{i,t}$ is the total quantity of purchases in the current month; $Active_{i,t}$ is a dummy variable that reflects the existence of any account activity in the current month; $BillExist_{i,t-1}$ is a dummy variable that reflects the existence of a bill with a non-zero balance in the previous balance; $Util_{i,t}$, for utilization, is debt divided by the credit limit; $\epsilon_{i,t}$ is an error term. Table 1 provides mnemonics, definitions and summary statistics for the independent variables used in our analyses.

Figure 1 plots the expected frequency of fees as a function of account tenure, holding the other control variables fixed at their means.[11] This analysis shows that fee payments are fairly common when accounts are initially opened, but that the frequency of fee payments declines rapidly as account tenure increases. In the first four years of account tenure, the monthly frequency of cash advance fee payments drops from 57% of all accounts to 13% of all accounts. The frequency of late fee payments drops from 36% to 8%. Finally, the frequency of over limit fee payments drops from 17% to 5%.[12]

Figures 2 reports the average value of each fee type as a function of account tenure. The

[10] The spline has knots every 12 months through month 72.

[11] Tenure in all figures starts at month two since borrowers cannot, by definition, pay late or over limit fees in the first month their accounts are open.

[12] We repeat the analysis controlling for behavior and FICO scores, both lagged by three months to reflect the fact that they are only computed quarterly. We find very similar patterns.

data plotted in Figure 2 is generated by estimating,

$$\begin{aligned} V_{i,,t}^{j} &= \alpha + \phi_i + \psi_{time} + Spline(Tenure_{i,t}) \\ &+ \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} \\ &\gamma_1 Util_{i,t-1} + \epsilon_{i,t}. \end{aligned}$$

The dependent variable $V_{i,t}^{j}$ is the *value* of fees of type j paid by account i at tenure t. All other variables are as before.

Figure 2 shows that, when an account is opened, the card holder pays \$6.65 per month in cash advance fees, \$5.63 per month in late fees, and \$2.46 per month in over limit fees. These numbers understate the total cost incurred by fee payments, as these numbers do not include interest payments on the cash advances, the effects of penalty pricing (i.e. higher interest rates), or the adverse effects of higher credit scores on other credit card fee structures. Like Figure 1, Figure 2 shows that the average value of fee payments declines rapidly with account tenure.

Figures 1 and 2 imply that fee payments substantially fall with experience. We next turn to a second pattern in our data.

2.2.2 The impact of past fee payment on current fee payment

There are four reasons to expect fee payments to have intertemporal linkages. First, the (cross-sectional) type of the card holder (e.g. forgetful) may influence fee paying behavior, causing fee payments to be positively autocorrelated; specifically, if person i pays a fee in period t then person i is more likely to be of the type that pays fees in general, implying that person i has a higher likelihood of paying a fee in period $t+k$ relative to other subjects in our sample. Second, transitory shocks that persist over more than one month (for instance, an unemployment spell) may influence fee paying behavior, causing fees to be positively autocorrelated. Third, transitory shocks that are negatively correlated across months (for instance, an annual vacation) will cause fees to be negatively autocorrelated. Fourth, fee payments may engender learning, causing fees to be negatively autocorrelated.

These four effects will jointly influence the following test statistic. We calculate:

$$L_{t,k} \equiv \frac{E\left[f_t \mid f_{t-k}=1\right]}{E\left[f_t\right]}$$

$$= \frac{\text{Probability of paying a fee given the agent paid a fee } k \text{ periods ago}}{\text{Probability of paying a fee}} \qquad (2)$$

To avoid Nickell bias, the expectations in $L_{t,k}$ are calculated *without* conditioning on the RHS variables from the previous subsection (most importantly, we do not use person fixed effects to calculate $L_{t,k}$). Specifically, $E\left[f_t\right]$ is just the average frequency of fee payments in period t.[13] Likewise, $E\left[f_t \mid f_{t-k}=1\right]$ is the average frequency of fee payments in period t among the subjects who paid a fee at time $t-k$.

Conditioning on this sparse information, a consumer who paid a fee k periods ago has a probability of paying a fee equal to the baseline probability, $E\left[f_t\right]$, multiplied by $L_{t,k}$. A value of 1 for $L_{t,k}$ indicates that having paid a fee k periods does not change the expected probability of paying a fee this period; a value less than one indicates lagged fee payment is associated with a reduction in the expected probability, and a value greater than one indicates lagged fee payment are associated with an increase in the expected probability. For example, if $L_{t,1}=0.6$, a consumer who paid a fee last month has a probability of paying a fee this month that is 40% below the baseline probability.

We report averages of $L_{t,k}$:

$$L_k \equiv \frac{1}{T}\sum_{t=1}^{T} L_{t,k}.$$

Hence, L_k is the average relative likelihood of paying a fee, if the account holder paid a fee k periods ago. The L_k statistic illustrates the some important time series properties in our data while avoiding econometric problems associated with estimating probit or logit models with fixed effects for a large N dataset. The statistic also avoids the fixed-effect biases in dynamic models (Nickell 1982).

Figure 3 plots L_k for all three types of credit card fees for values of k ranging from 1 to 35. All three lines start below 1, indicating that a fee payment last month is associated with a less than average likelihood of making a fee payment this month. For both cash advance and late fees, having paid a fee one month ago is associated with a 40% reduction in the

[13] Observations used for the calculation of L_k are from subjects who are in our sample at *both* date t and date $t-k$.

likelihood of paying a fee in the current month. For over limit fees, having paid a fee last month is associated with a 50% reduction in the likelihood of paying a fee in the current month.

The L_k plots rise with k, indicating that as a given fee payment recedes into the past, the negative association between this fixed lagged fee payment and current fee payments is diminished. By the time one year has passed, the association between the lagged fee payment and current fee payment has almost vanished.

For large values of k, all three graphs rise above 1. This asymptotic property probably reflects the variation in fee payments, which is driven by cross-sectional variation in the (persistent) type of the borrower. Some individuals have a relatively high long-run likelihood of paying fees. For instance, imagine that 20 percent of consumers never pay a fee (maybe because they are very disciplined), while the others have a long-run monthly probability $b > 0$ of paying a fee. Then, the long run L_k is $1/0.8 = 1.25$.[14]

In sum, the data show a robust time-series pattern. Paying a fee last month is associated with a sharply reduced likelihood of paying a fee this month. Paying a fee a year ago has little relationship to the likelihood of paying a fee now. Paying a fee two years ago is associated with a 20% elevation in the likelihood of paying a fee now.

Our findings imply that the there must be a mechanism that produces the short-run negative association. Moreover, this mechanism must be strong enough to temporarily overwhelm the positive long-run association in fee payments driven by type variation. We next present a model that captures these dynamic patterns.[15]

3 A simple model of learning and backsliding

The previous section showed that fee payments decline over time, and paying a past fee reduces the likelihood of paying a current fee, but with diminishing impact as the conditioned fee payment recedes into the past. In this section we describe a simple learning model, which can also be interpreted as a simple model of attention and inattention. This model includes three components: a stock of feedback, a dynamic updating equation, and a mapping from

[14] The numerator of (2) is b, while the denominator is $0.8b$. So $L_k = b/(0.8b) = 1/0.8$.

[15] The short-run drop in L_k would be even bigger if it were not offset by the positive autocorrelation in fees produced by both variation in types and transitory (multi-month) variation in fee-paying propensities.

the stock of feedback to the next month's fee payment.

Let F_t represent the effective stock of Feedback. Let $f_t \in \{0,1\}$ represent the current feedback. For simplicity, assume that experience is binary so that $f_t = 0$ (if you are not charged a fee at time t) and $f_t = 1$ (if you are). The stock of feedback, F_t, is:

$$F_t = \delta^t F_0 + \psi t + \sum_{s=1}^{t} \delta^{t-s} f_s. \tag{3}$$

In this dynamic updating equation, $\delta \in [0,1]$ represents the depreciation rate of the stock of feedback. If $\delta = 1$, there is no depreciation, and if $\delta = 0$, there is full depreciation after one period. We refer to this as "depreciation" but δ captures many related effects including recency bias, salience, forgetting, or any other form of temporal backsliding.[16] The term ψt means that, even if the consumer receives no feedback, she still learns with the passage of time, perhaps via word of mouth.

A consumer with stock of knowledge F_t will pay attention to the fees with probability $A(F_t)$, which we parametrize as:

$$\text{Probability of paying attention to the fee: } A(F_t) = C + De^{-\kappa F_t} \tag{4}$$

with $\kappa, C, D \geq 0$, and $C + D \leq 1$. We view Equation (4) as a representation of the psychology of attention, or memory. This equation implies that events that have happened relatively frequently are easier to remember. Also, after an extensive amount of learning, attention saturates to $C + D$. If $C + D = 1$, learning is perfect in the long run, but if $C + D < 1$, attention is imperfect, even in the long run.

A consumer faces two mutually exclusive types of opportunities. With probability p'', he can experience a need to pay a fee, for instance through financial stress. If that does not happen, he can just face an "avoidable opportunity" to pay a fee. If that opportunity happens, he pays of a fee if and only if he was not paying attention to fees. Symbolically:

$$\{f_{t+1} = 1\} = \{\{\text{Avoidable opportunity}\} \text{ and } \{\text{Don't pay attention}\}\} \text{ or } \{\text{Need}\}$$

Call p'' the probability of a need, p' the probability of an avoidable opportunity, and $P_{t+1} =$

[16] Rubin and Wenzel (1996) offer a comprehensive survey of the literature on forgetting.

$P_t(f_{t+1} = 1)$, the probability of paying a fee at $t + 1$, given the history up to time t. We have:

$$P_{t+1} = p'(1 - A(F_t)) + p''$$

and using the functional form (4), we get:

$$P_{t+1} = ae^{-\beta F_t/a} + b. \tag{5}$$

with $a = p'D$, $\beta = a\kappa$, $b = p'C + p''$. It is useful to define $\phi = \kappa\psi$, so the probability of paying a fee at $t + 1$ is:

$$P_{t+1} = a \exp\left(-\phi t - \frac{\beta}{a}\sum_{s=1}^{t}\delta^{t-s}f_s - \frac{\beta}{a}\delta^t F_0\right) + b \tag{6}$$

Controlling for person fixed effects, the more fees you have paid in the past, the more likely you are to avoid paying fees in the future: the first-order effect of a change in F_t on P_{t+1} is $-\beta$ (when the Taylor expansion is taken around $F_t = 0$). Parameter β captures the strength of short-term learning. When β is large, past feedback reduces the expected current rate of fee payment. Past fee payments drive down future fee payments through a learning mechanism, like reinforcement (see for example Camerer 2003 and Sutton and Barto 1998).

Under the above formulation, if the initial stock of knowledge $F_0 = 0$, the initial propensity to pay fees is $a + b$, while the long-run probability of a consumer who can pay perfect attention is b.

4 Results

4.1 Estimating the learning model

We estimate the model described in the previous section by the method of simulated moments (MSM). Define the vector of empirical moments—the L_k and tenure distribution discussed above—as $\bar{m}$. Given a parameter set θ, there is a theoretical moment set that we denote $m(\theta)$. Since we cannot directly calculate $m(\theta)$, we approximate it with a simulator, denoted $m_s(\theta, J_s)$, where J_s is the number of observations used in the simulation for m_s.

Pakes and Pollard (1989) show that

$$\hat{\theta} = \arg\min_{\theta} \left[(m_s(\theta, J_s) - \bar{m}) W (m_s(\theta, J_s) - \bar{m})' \right],$$

and this is a consistent estimator for θ_0, where W is an arbitrary positive definite weighting matrix. Intuitively, $\hat{\theta}$ is simply the parameter vector that minimizes the distance between the empirical moments $\bar{m}$ and the outputs of the structural model $m_s(\theta, J_s)$. In order to calculate the variance of $\hat{\theta}$, we need to define a few more terms. Let $M_\theta \equiv \partial m_s(\theta, J_s)/\partial\theta$ be the numerically calculated derivative of the simulated moments with respect to the parameter vector θ. We calculate the variance of $\bar{m}$ as

$$\Omega_M \equiv Var(m_s(\theta, J_m)), \tag{7}$$

where J_m is the number of observations from which $\bar{m}$ is calculated. Equation (7) defines the variance of $\bar{m}$ under the assumption that the model is correct. With these definitions in hand, following Pakes and Pollard (1989) and Laibson, Repetto and Tobacman (2007), we can calculate the variance of $\hat{\theta}$ as

$$Var(\hat{\theta}) = \left(M_\theta' W M_\theta\right)^{-1} M_\theta' W \left[\left(1 + \tfrac{J_m}{J_s}\right) \Omega_G \right] W M_\theta \left(M_\theta' W M_\theta\right)^{-1}.$$

We describe the simulation and optimization procedure in Appendix B.[17]

4.2 Estimation Results

Table 2 presents the MSM estimates for equation 6 for all three kinds of fees. The first row shows that rates of depreciation $1-\delta$ are quite substantial for all three fee types, ranging from nearly 20 percent for the late fee to 8 percent for the cash advance fee. Thus, the impact of paying a late fee a year ago has only about $.8^{11} \approx .085$ as much impact on the probability of paying a fee this month as having paid the same fee a month ago.

The second row provides estimates of β, which measures the rate of short-term learning. Higher values of β imply that the stock of feedback has a larger impact on the probability

[17] Using the method of simulated moments also allows us to avoid the downward bias in dynamic panel data models with fixed effects. We have also tried estimating such models, and found, within the limits of the bias, qualitatively similar results.

of paying a fee. Paying a fee this month reduces the probability of paying a fee next month by approximately a fraction β.[18]

The third row provides estimates of ϕ, which parameterizes the effect of the passage of time on fee payment. The effect on fee payment is small but present. For example, for the late fee, an additional month's time reduces fee probability by about 4%. The estimates also imply that in the first period, paying a fee once reduces the probability of learning by $\beta/(a\phi) = 21$ months.

The fourth through seventh rows give estimates for parameters a, b and F_0, which govern the long- and short-run properties of fee probability and provide an estimate of the initial stock of feedback. To allow for some of the heterogeneity in fee payment that may be responsible for L_k asymptoting to values above 1 for large values of k, we let b take two values, b^L and b^H, for the population. The initial stock of feedback F_0 is about equal to one fee payment, on average, for all borrowers. For $t = 0$, the initial probability of paying a fee is $a \exp\left(-\frac{\beta}{a}\delta^t F_0\right) + b$. The model estimates this to be about 34 to 37 percent for the late fee, about 10 to 12 percent for the over limit fee, and about 66 percent for the cash advance fee. The long-run propensity to pay a fee b is generally low for all types and all fees, ranging from 0 to 3 percent for the late fee, 0 to 2 percent for the over limit fee, and about 2 percent for the cash advance fee

The six panels of Figure 4 plot the actual moments against the estimated moments, thus giving some indication of goodness of fit. The model fits the tenure moments fairly closely for the late fee, although it predicts a somewhat more rapid initial decline in fee payments than seen in the data. The fit of the L_k is again good, although the model again predicts a more rapid decline in the impact of past fee payments and has a smaller asymptotic value. For the over limit fee, the estimated and actual moments nearly match for almost every value of k for the L_k graph. The tenure moments match somewhat more closely than in the late fees case, although the model still shows a more rapid reduction in fee payment than is seen in the data. The cash advance model fits less well. Although the tenure moments match to nearly the same degree as with the late fees, the L_k moments fail to match, with the model showing a relatively flat profile of moments. This last result may indicate that the determinants of cash advance fee payment are somewhat different from the other fees. This may not be surprising. Paying cash advance fees is less clearly a mistake than paying other

[18] As equation 6 shows, this is true when b is small, which is empirically the case.

kinds of fees, and may also be motivated by different kinds of liquidity needs than other kinds of fees (e.g. needing cash while traveling).

5 Discussion of alternative explanations not based on learning

We next discuss to what extent the patterns that we have observed are driven by other factors.

Potential correlation between financial distress and credit card tenure. The tendency to observe declining fees may reflect a tendency for new account holders to experience more financial/personal distress than account holders with high tenure. To test this hypothesis, we determined if FICO scores and Behavior scores (two inverse[19] measures of financial distress) correlate with account tenure.

Figure 5 plots FICO scores and behavior scores by account tenure, demeaned and normalized. To calculate the FICO variable, a single FICO mean is calculated for all accounts over all periods in our sample. This mean is used for the demeaning. A single FICO standard deviation is calculated for all accounts over all periods in our sample. This standard deviation is used for the normalization. An analogous method is used for the behavior score.

No time trend is apparent in the normalized data. To more formally test that the hypothesis, we predict FICO with an account-tenure spline using annual knots (controlling for account and time fixed effects). The estimated tenure spline exhibits slopes that bounce around in sign and are all very small in magnitude. For example, at a horizon of 5 years, the spline predicts a total (accumulated) change in the FICO score of 18 units since the account was opened. At a horizon of 10 years the spline predicts a total (accumulated) change in the FICO score of -0.04 units since the account was opened. Recall that the mean FICO score is 732 and the standard deviation of the FICO score is 81. Hence, financial distress does not show significant economic variation with account tenure.

[19] A high FICO or behavior score implies that the individual is a reliable creditor. Behavior scores are measures of credit risk internal to the financial institution designed to measure credit risk.

Potential correlation between purchasing patterns and credit card tenure. The tendency to observe declining fees may reflect a tendency for new account holders to spend more than account holders with high tenure. To test this hypothesis, we determined if purchases correlate with account tenure. Figure 5, which plots the demeaned and normalized level of purchases, again shows no economically significant time trend.

Non-utilization of the card. The fee dynamics that we observe could be driven by consumers who temporarily or permanently stop using the card after paying a fee on that card. We look for these effects by estimating a regression model in which the outcome of "no purchase in the current month" is predicted by dummies for past fee payments and control variables, including account and time fixed effects as well FICO, Behavior, and Utilization. We find very small effects of past fee payments on subsequent card use. For example, (controlling for account fixed effects) somebody who paid a fee every month for the past six months is predicted to be only 2% less likely to use their card in the next month relative to somebody with no fee payments in the last six months. Such very small effects cannot explain our learning dynamics, which are over an order of magnitude larger. Figure 5 also plots the absolute level of utilization (demeaned and normalized), exhibiting no time-series pattern).

Time-varying financial service needs. Time-varying financial service needs may also play an important role in driving service charge dynamics. To illustrate this idea, let ν_t represent a time-varying cost of time, so that

$$\Pr\left(f_t = 1\right) = \nu_t, \tag{8}$$

where ν_t is an exogenous process, that causes fee use, but is not caused by it. To explain our recency effect, one needs ν_t to be negatively autocorrelated at a monthly frequency. To see this, consider the regression,

$$f_t = \theta f_{t-1} + \text{controls}. \tag{9}$$

If (8) holds, then the regression coefficient is $\theta = cov\left(\nu_t, \nu_{t-1}\right)/var\left(f_{t-1}\right)$.

We run this regression, including all of our usual control variables, that is, time- and account-fixed effects, a tenure spline, *Purchase*, *Active*, *BillExist*, and *Util*. We also include *Behavior* and *FICO*.[20]

$$
\begin{aligned}
f_{i,t}^{j} &= \theta f_{i,t-1}^{j} + \alpha + \phi_i + \psi_{time} + Spline(Tenure_{i,t}) \\
&\quad + \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} \\
&\quad + \eta_4 FICO_{i,t-3} + \eta_5 Behave_{i,t-3} + \eta_6 Util_{i,t} + \epsilon_{i,t}.
\end{aligned}
$$

Results for the three types of fees are given in Table 3. We find that θ is -0.75 for the late fee, -0.52 for the over limit fee, and -0.28 for the cash advance fee. We call this the "recency effect," since the payment of a fee last month greatly reduces the probability that a fee will be paid this month.[21]

The empirical finding of $\theta < 0$ implies $corr\,(\nu_t, \nu_{t-1}) < 0$. Hence, to explain the "recency effect" with time-varying financial needs, it would need to be the case that ν_t is negatively autocorrelated. The autocorrelation of ν_t would need to be not only negative, but also greater than 0.75 (in the case of the late fee) in absolute value: $corr\,(\nu_t, v_{t-1}) \leq \theta = -0.75$.[22]

We think that such a very strong negative autocorrelation of monthly needs is very unlikely.[23] First, since the regression results include time fixed effects, such autocorrelations could not occur from events that happen at regular intervals during the year — e.g., from summer vacations. Second, the presence of highly negative autocorrelations at a monthly level would rule out events that last more than one month. For example, a personal crisis that raised the opportunity cost of time for two months would create a positive autocorrelation in time needs and fee payments over the two months, not a negative one. Third, the time-varying needs would have to produce a higher than average fee payment in one month

[20] The results do not differ if we instead begin the regressions in month 2 and exclude the Behavior and FICO scores.

[21] There is a potential small sample bias (Nickell 1981), to which we thank Peter Fishman and Devin Pope for drawing our attention. To see how large it is, we note that if f_t is i.i.d., then in the regression $f_t = \theta f_{t-1}$+constant, done over a T periods, the expected value of θ is $-1/T$. With $T = 24$, the bias is -0.05. We conclude that, in our study, the small sample bias is very small compared to the large negative θ that we find.

[22] It is easy to see that under (8), $cov\,(f_t, f_{t-1}) = cov\,(\nu_t, \nu_{t-1})$, and $var\,(f_t) = E\,[\nu_t]\,(1 - E\,[\nu_t]) \geq E\left[\nu_t^2\right] - E\,[\nu_t]^2 = var\,(\nu_t)$, as $\nu_t \in [0,1]$. So, $\theta = cov\,(f_t, f_{t-1})\,/var\,(f_{t-1})$ satisfies $|\theta| \leq |cov\,(\nu_t, \nu_{t-1})|\,/var\,(\nu_t) = |corr\,(\nu_t, \nu_{t-1})|$, and θ and $corr\,(\nu_t, \nu_{t-1})$ have the same sign.

[23] The least implausible type of negatively autocorrelated process in economics is a "periodic spike" process, which take a value of a every K periods, and $b \neq a$ otherwise. It has an autocorrelation of $-1/\,(K-1)$. We fail to find evidence for such a pattern in credit card use other than fees. For instance, expenses across time are *positively* autocorrelated.

followed by a lower than average fee payment in the following month. This would rule out episodes of high opportunity cost of time for one month followed by a return to the status quo.

For most plausible processes, needs are likely to be positively autocorrelated. For example, the available evidence implies that income processes are positively autocorrelated.[24] While we cannot rule out the "negatively autocorrelated needs" story, existing microeconomic evidence suggests it is highly unlikely to be the right explanation for the empirical patterns that we observe. We conclude that the finding of $\theta < 0$ in (9) is most plausibly explained by a recency effect – consumers become temporarily vigilant about fee avoidance immediately after paying a fee.

6 Conclusion

Credit card users learn about add-on fees by paying them. With years of experience, credit card customers substantially reduce these fee payments. We document this process using a three-year panel dataset representing 120,000 accounts.

In our data, new accounts generate direct fee payments of $15 per month. The data implies that negative feedback — i.e., paying fees — teaches consumers to avoid triggering fees in the future. Controlling for account fixed effects, monthly fee payments fall by 75% during the first four years of account life.

We also find that learning is not monotonic. In our basic specification, we estimate that knowledge depreciates 10% per month. As previous fee-paying lessons recede into the past, consumers tend to backslide. However, on net, knowledge accumulation dominates knowledge depreciation. Over time, fee payments drastically fall.

[24] See, for example, Guvenen (2007).

References

Agarwal, Sumit, Souphala Chomsisengphet, Chunlin Liu and Nicholas S. Souleles (2007). "Do Consumers Choose the Right Credit Contracts?" Manuscript, The Wharton School, University of Pennsylvania.

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix and David I. Laibson (2007). "The Age of Reason: Financial Mistakes Over the Lifecycle." NBER Working Paper 13191.

Agarwal, Sumit, Chunlin Liu and Nicholas S. Souleles (forthcoming). "The Reaction of Consumer Spending and Debt to Tax Rebates– Evidence from Consumer Credit Data." *Journal of Political Economy*.

Ameriks, John, Andrew Caplin, and John Leahy (2004). "The Absent-Minded Consumer". NBER Working Paper 10216.

Anderson, J. R. (2000). *Learning and Memory: An Integrated Approach*, 2nd ed. Wiley, New York.

Argote, Linda, Sara Beckman and Dennis Epple (1990). "The Persistence and Transfer of Learning in Industrial Settings." *Management Science*, 36(1), 141-151

Aumann, Robert, Sergiu Hart and Motty Perry (1997). "The Absent-Minded Driver." *Games and Economic Behavior*, 20(1), 102-116.

Ausubel, Lawrence M. (1991). "The Failure of Competition in the Credit Card Market." *American Economic Review*, 81(1), 50-81.

———, (1999). "Adverse Selection in the Credit Card Market." Manuscript, University of Maryland.

Bahk, B. and M. Gort (1993). "Decomposing Learning by Doing in New Plants." *Journal of Political Economy*, 101(4), 561-583.

Bailey, Charles. (1989). "Forgetting and the Learning Curve: A Laboratory Study." *Management Science*, 35(2), 340-352.

Barber, Brad M., Terrance Odean and Michal Strahilevitz (2004). "Once Burned, Twice Shy: naïve Learning, Counterfactuals,and the Repurchase of Stocks Previously Sold." Manuscript, UC Davis.

Barrios, S. and E. Strobl (2004). "Learning by Doing and Spillovers: Evidence from Firm-Level Panel Data." *Review of Industrial Organization*,25(1), 175-203.

Bénabou, Roland and Jean Tirole (2002). "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics*, 117(4), 871-913.

Benkard, C.Lanier (2000). "Learning and Forgetting: The Dynamics of Aircraft Production." *American Economic Review*, 90(6), 1034-1054.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman (2007). "What's Advertising Content Worth? A Field Experiment in the Consumer Credit Market." Manuscript, University of Chicago Graduate School of Business.

Calem, Paul S. and Loretta J. Mester (1995). "Consumer Behavior and the Stickiness of Credit-Card Interest Rates." *American Economic Review*, 85(6), 1327-1336.

Calem, Paul S., Michael B. Gordy and Loretta J. Mester (200x). "Switching Costs and Adverse Selection in the Market for Credit Cards." *Journal of Banking & Finance*, 30(6), June 2006, 1653-1685 .

Camerer, Colin F. (2003). *Behavioral Game Theory.* Russell Sage Foundation and Princeton University Press, Princeton, NJ.

Crawford, Vincent P. (1995). "Adaptive Dynamics in Coordination Games." *Econometrica*, 63(1), 103-143.

Darr, Eric, Linda Argote and Dennis Epple (1995). "The Acquisition, Transfer, and Depreciation of Knowledge in Service Organizations: Productivity in Franchises." *Management Science*, 41( ), 1750-1762.

DellaVigna, Stefano, and Ulrike Malmendier (2005), "Paying Not To Go To The Gym," *American Economic Review*, 96(3), 694-719, DellaVigna, Stefano and Ulrike Malmendier (2004). "Contract Theory and Self-Control: Theory and Evidence." *Quarterly Journal of Economics*, 119(2), 353-402.

Evans, George W. and Seppo Honkapohja (2001). *Learning and Expectations in Macroeconomics.* Princeton University Press, Princeton, NJ.

Evans, George W., Seppo Honkaphoja and Ramon Marimon (2000). "Convergence in Monetary Inflation Models with Heterogeneous Learning Rules." *Macroeconomic Dynamics*, 5(1), 1-31.

Fishman, Peter and Devin G. Pope (2006), "Specific Deterrence: Evidence and Implications," Berkeley Working Paper.

Gabaix, Xavier and David I Laibson (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, 121(2), May 2006, p. 505-540.

Gross, David B. and Nicholas S. Souleles (2002a). "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data." *Quarterly Journal of Economics*, 117:1, 149-185.

——— (2002b). "An Empirical Analysis of Personal Bankruptcy and Delinquency." *The Review of Financial Studies*,15(1), 319-347.

Gruber, H (200x). "The Learning Curve in the Production of Semiconductor Memory Chips." *Applied Economics*, 24(8), 885-894.

Guvenen, Fatih (2007). "Learning Your Earnings: Are Labor Income Shocks Really Very Persistent?" *American Economic Review*, 97(3), 687-712.

Harrison, Glenn and John A. List. "Field Experiments," *Journal of Economic Literature* (2004), 42 (December): 1013-1059.

Heidhues, Paul, and Botond Koszegi (2005), "Loss Aversion, Price Stability, and Sales," Manuscript, University of California, Berkeley.

Ho, Teck-Hua and Juin-Kuan Chong (2003). "A Parsimonious Model of Stockkeeping-Unit Choice." *Journal of Marketing Research*, 40(August), 351-365.

Jin, Ginger, and Philip Leslie, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, 118 (2003), 409–51.

Kerr, Sougata and Lucia Dunn (2002). "Consumer Search Behavior in the Changing Credit Card Market." Manuscript, Ohio State University.

Kerr, Sougata, Stephen Cosslett, and Lucia Dunn (2004). "Do Banks Use Private Information from Consumer Accounts? Evidence of Relationship Lending in Credit Card Interest Rate Heterogeneity."Manuscript, Ohio State University.

Koszegi, Botond, and Matthew Rabin (2006), "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121(4), pp. 1133-1166

Laibson, David, Andrea Repetto and Jeremy Tobacman (2007). "Estimating Discount Functions with Consumption Choices over the Lifecycle," NBER Working Paper 13314.

Lehrer, Ehud (1988). "Repeated Games with Stationary Bounded Recall Strategies." *Journal of Economic Theory*, 46(1), 130-144.

Lemieux, Thomas and W. Bentley MacLeod (2000). "Supply Side Hysteresis: The Case of the Canadian Unemployment Insurance System." *Journal of Public Economics*,78(1), 139-170.

List, John A. "Does Market Experience Eliminate Market Anomalies?," *Quarterly Journal of Economics* (2003), 118(1), 41-71.

Lucas, R.E. (1986). "Adaptive Behavior and Economic Theory." In R.M. Hogarth & M.W. Reder (eds.). *Rational Choice. Chicago: University of Chicago Press.* 217-242

McAfee, R. Preston and John McMillan (1996). "Analyzing the Airwaves Auction." *Journal of Economic Perspectives*, 10(1), 159-175.

McAllister, Patrick H. (1991). "Adaptive Approaches to Stochastic Programming." *Annals of Operations Research*, 30(1), 45-62.

McGaugh, James (2000). "Memory–A Century of Consolidation." *Science*, 287( ), 248-251.

Malmendier, Ulrike, and Devin Shanthikumar (2005), "Are Small Investors naïve about Incentives?" *Journal of Financial Economics*, August 2007, 85:2, 457-489.

Marimon, Ramon and S. Sunder (1994). "Expectations and Learning under Alternative Monetary Regimes: An Experimental Approach," *Economic Theory* 4(1), 131-162.

Massoud, Nadia, Anthony Saunders and Barry Scholnick (2006). "The Cost of Being Late: The Case of Credit Card Penalty Fees." mimeo, University of Alberta and New York University.

Miao, Chun-Hui (2005), "Consumer Myopia, Standardization, and Aftermarket Monopolization," mimeo, University of South Carolina.

Miravete, E. J. (2003). "Choosing the Wrong Calling Plan? Ignorance and Learning." *American Economic Review*, 93(2), 297-310.

Mullainathan, Sendhil (2002). "A Memory-Based Model of Bounded Rationality." *Quarterly Journal of Economics*, 117: 735-773.

Mullainathan, Sendhil, and Andrei Shleifer (2005), "The Market for News," *American Economic Review,* 95, 1031-1053.

Nickell, Stephen (1981) "Biases in Dynamic Models with Fixed Effects.". *Econometrica*, Vol. 49(6), 1417-1426.

Nye, W. (1996). "Firm-Specific Learning-by-Doing in Semiconductor Production: Some Evidence from the 1986 Trade Agreement." *Review of Industrial Organization*, 11(3), 383-394.

Oster, Sharon M., and Fiona M. Scott Morton (2005), "Behavioral Biases Meet the Market: The Case of Magazine Subscription Prices," *Advances in Economic Analysis and Policy*, 5 ().

Pakes, Ariel M. and Jonathan Pollard (1989), "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57(5), pages 1027-57, September.

Piccione, Michele and Ariel Rubinstein (1997a). "On the Interpretation of Decision Problems with Imperfect Recall." *Games and Economic Behavior*, 20(1), 3-24.

Ravina, Enrichetta (2005). "Habit Persistence and Keeping Up with the Joneses: Evidence from Micro Data." Manuscript, New York University.

Roth, Alvin E. and Ido Erev (1995). "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term." *Games and Economic Behavior*, 8(2), 164-212.

Rubin, David and Amy Wenzel (1996). "One Hundred Years of Forgetting: A Quantitative Description of Retention." *Psychology Review*, 103: 734-760.

Sargent, Thomas J (1999). *The Conquest of American Inflation.* Princeton University Press, Princeton, NJ.

Shui, Haiyan and Lawrence M. Ausubel (1995). "Time Inconsistency in the Credit Card Market."Manuscript, University of Maryland.

Spiegler, Rani (2006), "The Market for Quacks," *Review of Economic Studies*, 2006.

Sutton, Richard S. and Andrew G. Barto (1998), *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA.

Thompson, Peter (2001). "How Much Did the Liberty Shipbuilders Learn? New Evidence for an Old Case Study." *Journal of Political Economy*, 109(1), 103-137.

Thornton, R. and P. Thompson (2001). "Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuidling." *American Economic Review*, 91(6), 1350-1368.

Van Huyck, John B., Raymond C. Battalio and R. Beil (1990). "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." *American Economic Review*, 80(2), 234-248.

Van Huyck, John B., Raymond C. Battalio and R. Beil (1991). "Strategic Uncertainty, Equilibrium Selection Principles, and Coordination Failure in Average Opinion Games." *Quarterly Journal of Economics*, 106(6), 885-910.

Van Huyck, John B., Raymond C. Battalio and Frederick W. Rankin (2001). "Selection Dynamics and Adaptive Behavior Without Much Information." Manuscript, Texas A&M University.

Van Huyck, John B., J. P. Cook and Raymond C. Battalio (1994). "Selection Dynamics, Asymptotic Stability, and Adaptive Behaviour."*Journal of Political Economy*, 102(6), 975-2005.

Wixted, John (2004a). "One Hundred Years of Forgetting: A Quantitative Description of Retention." *Psychology Review*, 111, 864-879.

——— (2004b). "The Psychology and Neuroscience of Forgetting." *Annual Review of Psychology*, 55, 235-269.

Zimmerman, M. (1982). "Learning Effects and the Commercialization of New Energy Technologies: The Case of Nuclear Power." *Bell Journal of Economics*, 13(2), 297-310.

Appendix A: Data Description

The total sample consists of 125,384 accounts open as of January 2002 and 22,392 opened between January and December of 2002 observed through December 2004. These accounts were randomly sampled from several million accounts held by the bank. From this sample of 147,776, we drop accounts that were stolen, lost, or frozen (due to fraud). We also exclude accounts that do not have any activity (purchases and payments) over the entire period. This leaves 128,142 accounts. Finally, we also remove account observations subsequent to default or bankruptcy, as borrowers do not have the opportunity to pay fees in such instances. This leaves us with an unbalanced panel with 3.9 million account observations.

Table A1 provides summary statistics for variables related to the accounts, including account characteristics, card usage, fee payment, and account holder characteristics. The second column notes whether the variable is observed monthly ('M'), quarterly ('Q'), or at account origination ('O'), the third column reports variable means, and the fourth column variable standard deviations. Note that the monthly averages for the 'Fee Payment' variables imply annual average total fees paid of \$141 (=\$11.75×12), with about 7.52 fee payments per year. Higher interest payments induced by paying fees (which raise the interest rate on purchases and cash advances) average about \$226 per year.

The accounts also differ by how long they have been open. Over 31 percent of the accounts are less than 12 months old, 20 percent are between 12 and 24 months old, 18 percent are between 24 and 36 months old, 13 percent are between 36 and 48 months old, 10 percent are between 48 and 60 months old, and 8 percent are more than 60 months old.

Table A1: Variable Descriptions and Summary Statistics

Description (Units)	Freq.	Mean	Std. Dev.
Account Characteristics			
Interest Rate on Purchases	M	14.40	2.44
Interest Rate on Cash Advances (%)	M	16.16	2.22
Credit Limit ($)	M	8,205	3,385
Card Usage			
Current Cash Advance ($)	M	148	648
Payment ($)	M	317	952
New Purchases ($)	M	303	531
Debt on Last Statement ($)	M	1,735	1,978
Minimum Payment Due ($)	M	35	52
Utilization (Debt/Limit) (%)	M	29	36
Fee Payment			
Total Fees ($)	M	10.10	14.82
Cash Advance Fee ($)	M	5.09	11.29
Late Payment Fee ($)	M	4.07	3.22
Over Limit Fee ($)	M	1.23	1.57
Extra Interest Payments:			
... Due to Over Limit or Late Fee ($)	M	15.58	23.66
... Due to Cash Advances ($)	M	3.25	3.92
Number of Times per month			
... Cash Advance Fee Paid	M	0.38	0.28
... Late Fee Paid	M	0.14	0.21
... Over Limit Fee Paid	M	0.08	0.10
Borrower Characteristics			
FICO (Credit Bureau Risk) Score	Q	731	76
Behavior Score	Q	727	81
Number of Credit Cards	O	4.84	3.56
Number of Active Cards	O	2.69	2.34
Total Credit Card Balance ($)	O	15,110	13,043
Mortgage Balance ($)	O	47,968	84,617
Age (Years)	O	42.40	15.04
Income ($)	O	57,121	114,375

Notes: The "Credit Bureau Risk Score" is provided by Fair, Isaac and Company (hence 'FICO'). The greater the score, the less risky the consumer is. The "Payment Behavior Score" is a proprietary score based on the consumer's past payment history and debt burden, among other variables. It is created by the bank to capture determinants of consumer payment behavior not accounted for by the FICO score. "Q" indicates the variable is observed quarterly, "M" monthly, and "O" only at account origination.

Appendix B: Simulation and Optimization Procedure

The simulation procedure was created to match the data sample as closely as possible. The data consist of approximately 120,000 accounts observed over 35 periods. Accounts have tenure ranging from 1-72 months. In order to calculate the tenure moments, we simulate a group of 120,000$\times\frac{36}{72}$ agents over 59 periods.[25]

The L_k are estimated with a simulation of a separate sample of agents. Because the estimated L_k vary depending on the distribution of tenure in the sample, we need to take into account entry and exit from the sample population.[26] In each period, $N/36$ agents enter the simulation with tenure of 0 months, where $N = J_S \times 36$ is the total number of observations.

Given a set of agents each with their own stock of knowledge, the simulation proceeds as follows. First, we calculate the probability of each agent paying a fee. Random numbers are then drawn for each agent from a uniform distribution on the interval $[0, 1]$. If an agent's draw is less than her probability of a fee payment, then she is recorded as paying a fee in this period. Next, a random 1 percent of the agents are removed from the sample, corresponding to the rate of attrition observed in the data. Any agent with tenure of 72 months is also removed from the sample. In the final step, we update each agent's knowledge stock according to whether or not she paid a fee, and we add a new set of $N/36$ agents with zero tenure. The L_k simulation is run for 108 periods. The first 72 generate a full distribution of agents

[25] 120,000*35 is the total number, N, of person*month observations. N*$\frac{59}{72}$ is approximately the number of these observations that are used in the regression sample to calculate the tenure moments from the data. The group of people that is simulated over the 59 periods therefore numbers N/72.

[26] Since the change in an agent's probability of paying a fee next period given this period is dependent on her knowledge stock, and the magnitude of her knowledge stock is correlated with her tenure, her response to a fee event will be correlated with her tenure.

with tenure between 1 and 72 months. Periods 73-108 are the sample from which the L_k are calculated. This calculation proceeds identically to that used on the empirical data set.

The optimization procedure uses a combination of a grid search and a hill-climbing algorithm. The first stage employs a coarse grid that searches over a broad set of possible parameter values using a relatively low value of J_s, for example, 40,000. The grid search is split into a number of separate samples. Consider a grid consisting of $A \times B$ points. The grid search calculates the value of the objective function at each point. In practice, this is performed with multiple programs; for example, there could be A programs that each calculate one $1 \times B$ column of grid points. The best point from each group is then saved. If any of the top 5 points is on or near the edge of the grid, we expand the grid around that point to ensure we find the global maximum. In practice, more than 40 search programs were used, generating 40+ candidate points from which to run the hill-climber.

The second stage of the optimization employs MATLAB's fminsearch command, which carries out the Nelder-Mead simplex method—the hill-climber. The optimization is run until it converges to a point such that the estimated parameters do not change for at least 20 iterations. This generally requires at least 120 total iterations. J_s was set to 180,000 to insure that the point to which the algorithm converged was not an abnormally low value due to measurement error.[27]

In order to calculate Ω_M, we run the simulation 1000 times using J_M agents at the optimal $\hat{\theta}$ and then calculate the variance-covariance matrix of the 1000 sets of simulated moments. To calculate M_θ we perturb each parameter individually by ± 1 percent. We then calculate numerical derivatives for the moments with respect to the parameter vector, creating, in the end, two 94×7 matrices of derivatives; one from the positive perturbations, one from the negative perturbations. We choose the smaller of the two estimates of the derivative for each component of the moment vector to ensure that we have an upper bound for $Var(\hat{\theta})$.

[27] Intuitively, near the optimum, the objective function gets flat. As the slope of the objective function decreases, differences in its measured value are more likely to be due to simulation error than true variation.

Table 1: Regression Variable Mnemonics and Summary Statistics

Mnemonic	Description	Mean	Std. Dev.
$f_{i,t}^{Late}$	Dummy for Late Fee Payment at t	0.18	0.21
$f_{i,t}^{Over}$	Dummy for Over Limit Fee Payment at t	0.08	0.10
$f_{i,t}^{Advance}$	Dummy for Cash Advance Fee Payment at t	0.38	0.28
$F^{Late}(\delta)_{i,t-1}$	Net Times Late Fees Paid Through $t-1$	1.58	0.94
$F^{Over}(\delta)_{i,t-1}$	Net Times Over Limit Fees Paid Through $t-1$	0.74	0.42
$F^{Advance}(\delta)_{i,t-1}$	Net Times Cash Advance Fees Paid Through $t-1$	2.38	0.78
$FICO_{i,t-3}$	FICO Score	727	81
$Behave_{i,t-3}$	Behavior Score	731	76
$Util_{i,t}$	Utilization (Debt/Limit)	29	36
$Purchase_{i,t}$	Purchases ($) at t	303	531
$Active_{i,t}$	Dummy for Account Activity (Purchases) at t	0.86	0.19
$BillExist_{i,t-1}$	Dummy for Existence of a Bill at $t-1$	0.82	0.15

Notes: $f_{i,t}^{j}$ denotes a payment of fee type j by account i at tenure t. $F^j(\delta)_{i,t-1} = f_{i,t-1}^{j} + (1-\delta)F^j(\delta)_{i,t-2}$, i.e. the total number of fees of type j paid through tenure $t-1$, less those forgotten. Means and standard deviations for these variables are computed for the (three) values of δ estimated below in Table 3. $FICO$ denotes the credit bureau risk score provided by Fair, Isaac and Company, lagged one quarter. The greater the score, the less risky the consumer is. The Behavior score is a propriety number created by the bank to capture determinants of consumer payment behavior not accounted for by the FICO score. Utilization is the ratio of current debt on the account to the current credit limit. Purchases is the dollar amount of purchases on the account, while the dummy variable for account activity is one if there were any purchases on the account. The bill existence dummy is one if the consumer received a bill at tenure $t-1$.

Table 2: Model Estimation Results

	Late	Over Limit	Cash Advance
δ	0.8007	0.9187	0.9060
	(0.0016)	(0.0020)	(0.0002)
β	0.8547	0.9506	0.2157
	(0.0041)	(0.0067)	(0.0007)
ϕ	0.0435	0.0452	0.0282
	(0.0003)	(0.0005)	(0.0007)
a	0.9040	0.5438	0.8410
	(0.0026)	(0.0045)	(0.0009)
b^L	0.0000	0.0003	0.0213
	(0.0016)	(0.0024)	(0.0001)
b^H	0.0298	0.0198	0.0195
	(0.0002)	(0.0001)	(0.0001)
$\overline{F_0}$	1.0227	0.9746	1.0317
	(0.0068)	(0.0064)	(0.0068)

Notes: This table presents MSM estimates of (6), $P_t\left(f_{t+1}=1\right) = a \exp\left(-\phi t - \frac{\beta}{a}\sum_{s=1}^{t}\delta^{t-s}f_s - \frac{\beta}{a}\delta^t F_0\right) + b$ for each fee. We allow for two values of b, b^L and b^H to permit different long-run propensities to pay fees within the population. Details of the simulations are given in Appendix B. Standard errors, in parentheses, are computed via Monte Carlo methods.

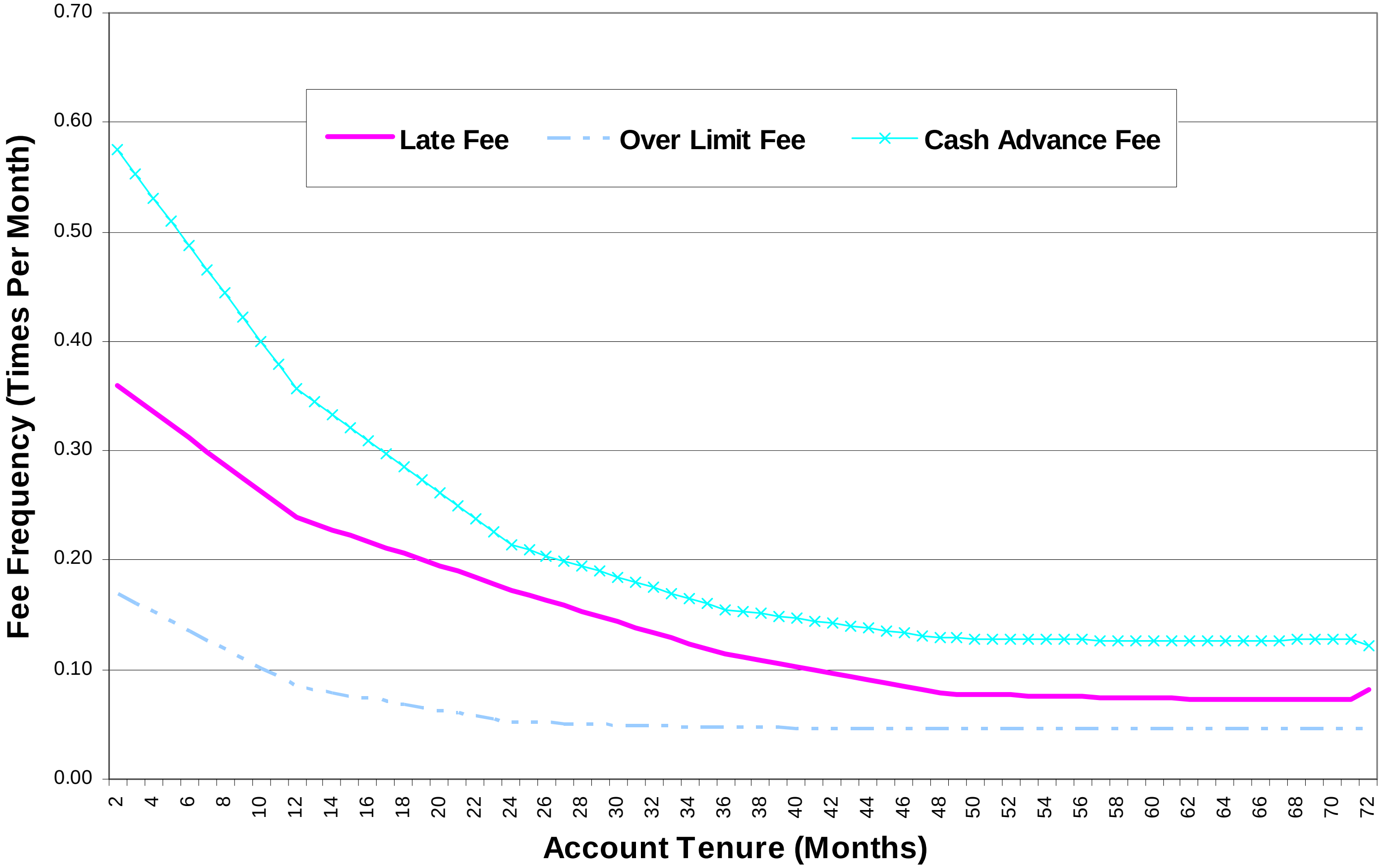
Table 3: Time-Varying Needs

	$f_{i,t}^{Late}$	$f_{i,t}^{Over}$	$f_{i,t}^{Advance}$
Intercept	0.2487*	0.1267**	0.4263**
	(0.08236)	(0.0284)	(0.1547)
$f_{i,t-1}^{j}$	-0.7483**	-0.5248**	-0.2784**
	(0.1403)	(0.1076)	(0.0640)
$t <= 12$	-0.0103*	-0.0077	-0.0237**
	(0.0037)	(0.0024)	(0.0068)
$12 < t <= 24$	-0.0059*	-0.0025**	-0.0127**
	(0.0023)	(0.0010)	(0.0048)
$24 < t <= 36$	-0.0044*	-0.0004	-0.0057*
	(0.0019)	(0.0001)	(0.0028)
$36 < t <= 48$	-0.0021	-0.0001	-0.0021
	(0.0015)	(0.0001)	(0.0018)
$48 < t <= 60$	-0.0003	-0.0001	-0.0002
	(0.0016)	(0.0001)	(0.0068)
$60 < t$	-0.0001	-0.0001	-0.0001
	(0.0014)	(0.0001)	(0.0008)
$Purchase/100_{i,t}$	0.0052	0.0021**	0.0073
	(0.0034)	(0.0003)	(0.0053)
$Active_{i,t}$	0.0071	0.0026**	0.0093
	(0.0048)	(0.0009)	(0.0058)
$BillExist_{i,t-1}$	0.0618**	0.0179*	0.0964**
	(0.0257)	(0.0084)	(0.0389)
$Behave_{i,t-3}$	-0.0035**	-0.0028**	-0.0053*
	(0.0008)	(0.0007)	(0.0025)
$FICO_{i,t-3}$	-0.0027**	-0.0014**	-0.0046*
	(0.0005)	(0.0006)	(0.0021)
$Util_{i,t}$	0.0506**	0.0283**	0.0693**
	(0.0074)	(0.008)	(0.0182)
Adjusted R-squared	0.0416	0.0484	0.0497
No. of Obs.	3.9 million	3.9 million	3.9 million

Notes: This table reports the results of estimating $f^j_{i,t} = \alpha + \phi_i + \psi_{time} + \theta f^j_{i,t-1} + Spline(Tenure_{i,t}) + \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} + \eta_4 FICO_{i,t-3} + \eta_5 Behave_{i,t-3} + \eta_6 Util_{i,t} + \epsilon_{i,t}$, where the first three terms are a constant, and account- and time- fixed effects. Rows 3 through 8 report the coefficients on the spline for account tenure (where the spline has yearly knot points). Variable definitions are as in table 1. Huber/White/Sandwich standard errors are in parentheses.

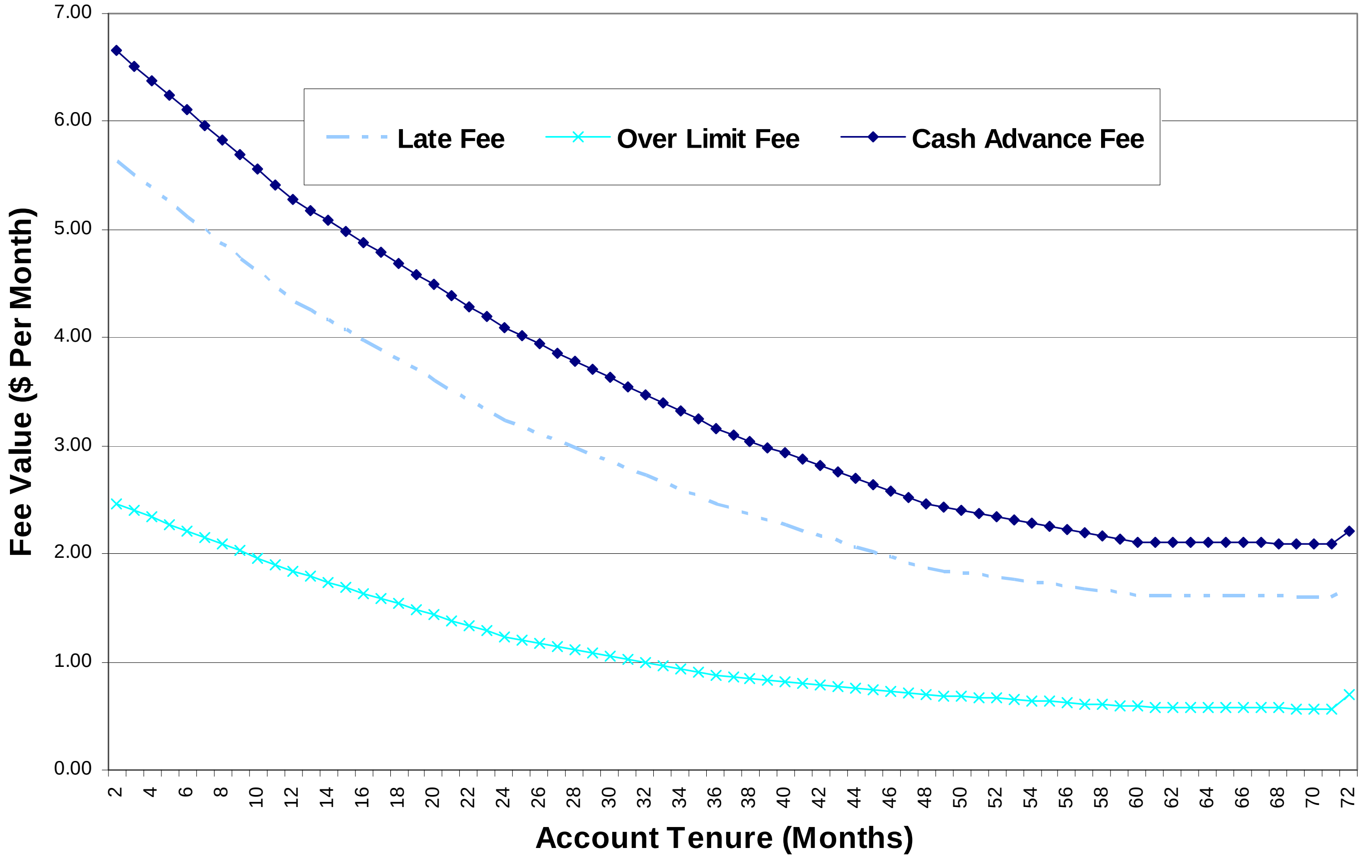
* denotes statistical significance at a 95 percent confidence level, and ** denotes statistical significance at a 99 percent confidence level.

Figure 1: Fee Frequency and Account Tenure



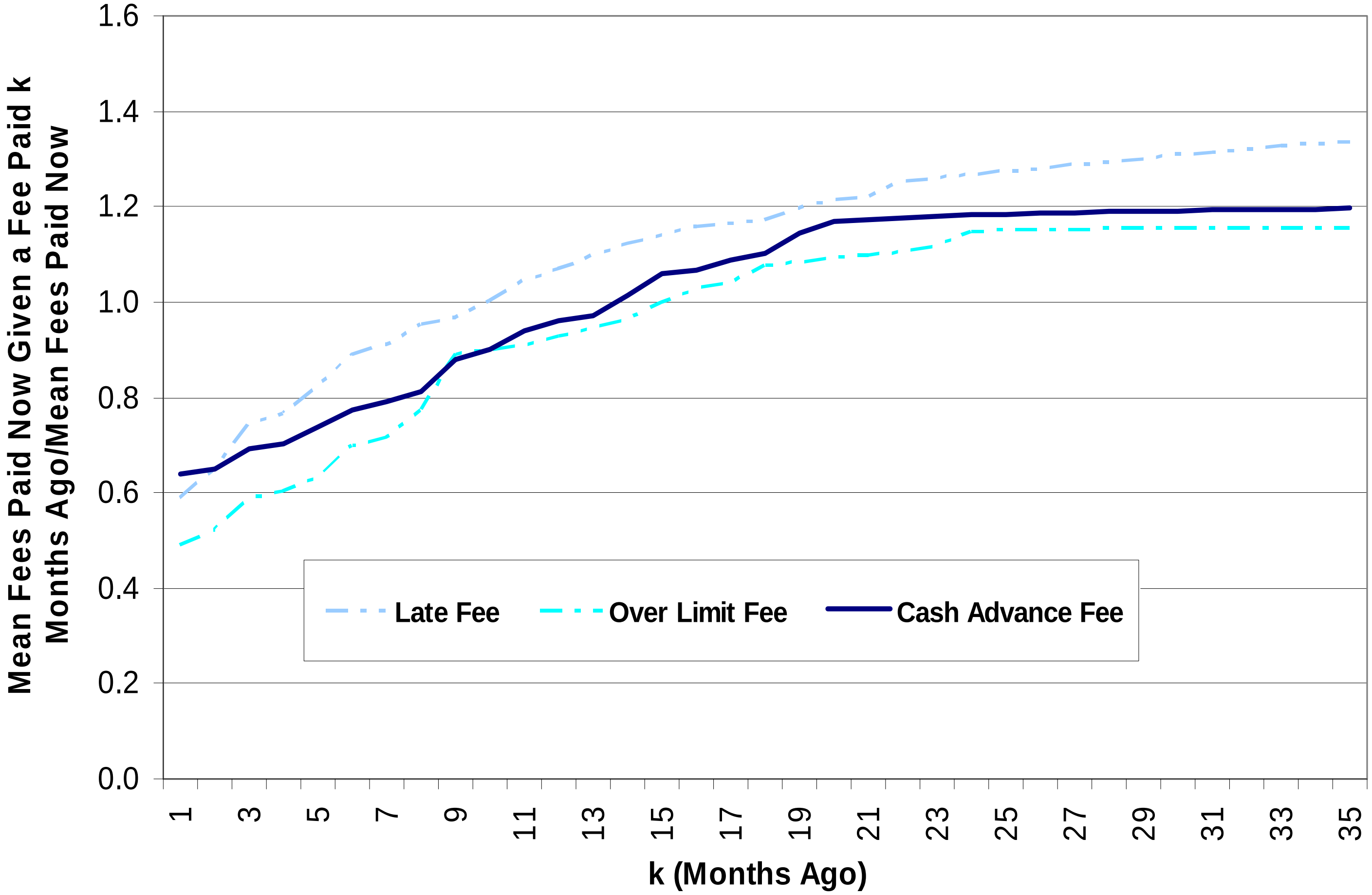
Notes: This figure plots the fitted values of regressions of fee frequency (times per month fees are paid) on a continuous piecewise linear function of account tenure (the function is a spline, with knots every twelve months, on the time since the account was opened), a constant, account- and time-fixed effects, and control variables (utilization (debit/limit), purchase amount, and dummy variables for any account activity this month and the existence of a bill last month). The intercept is computed by summing the constant with the product of the estimated coefficients on the control variables and their average values (the account and time-fixed effects sum to zero by construction). Tenure starts at the second month because account holders are, by definition, unable to pay late or over limit fees in their first month of account tenure.

Figure 2: Fee Value and Account Tenure



Notes: This figure plots the fitted values of regressions of fee value (dollars per month in fees paid) on a continuous piecewise linear function of account tenure (the function is a spline, with knots every twelve months, on the time since the account was opened), a constant, account- and time-fixed effects, and control variables (utilization (debit/limit), purchase amount, and dummy variables for any account activity this month and the existence of a bill last month). The intercept is computed by summing the constant with the product of the estimated coefficients on the control variables and their average values (the account and time-fixed effects sum to zero by construction). Tenure starts at the second month because account holders are, by definition, unable to pay late or over limit fees in their first month of account tenure.

Figure 3: Impact of Fees Paid k Months Ago on Fees Paid Now



Notes: This figure plots $L_k=E(f_t|f_{t-k}=1)/E(f_t)$, the ratio of the conditional mean of fees f_t paid now given a fee was paid k months ago to the mean of fees paid now. If this value is 1, having paid a fee k months ago has no effect on current fee payment; if it is less than one, having paid a fee k months ago reduces current fee payment; if it is greater than one, it increases fee payment.

Figure 4: Actual and Predicted Moments for Tenure and L_k

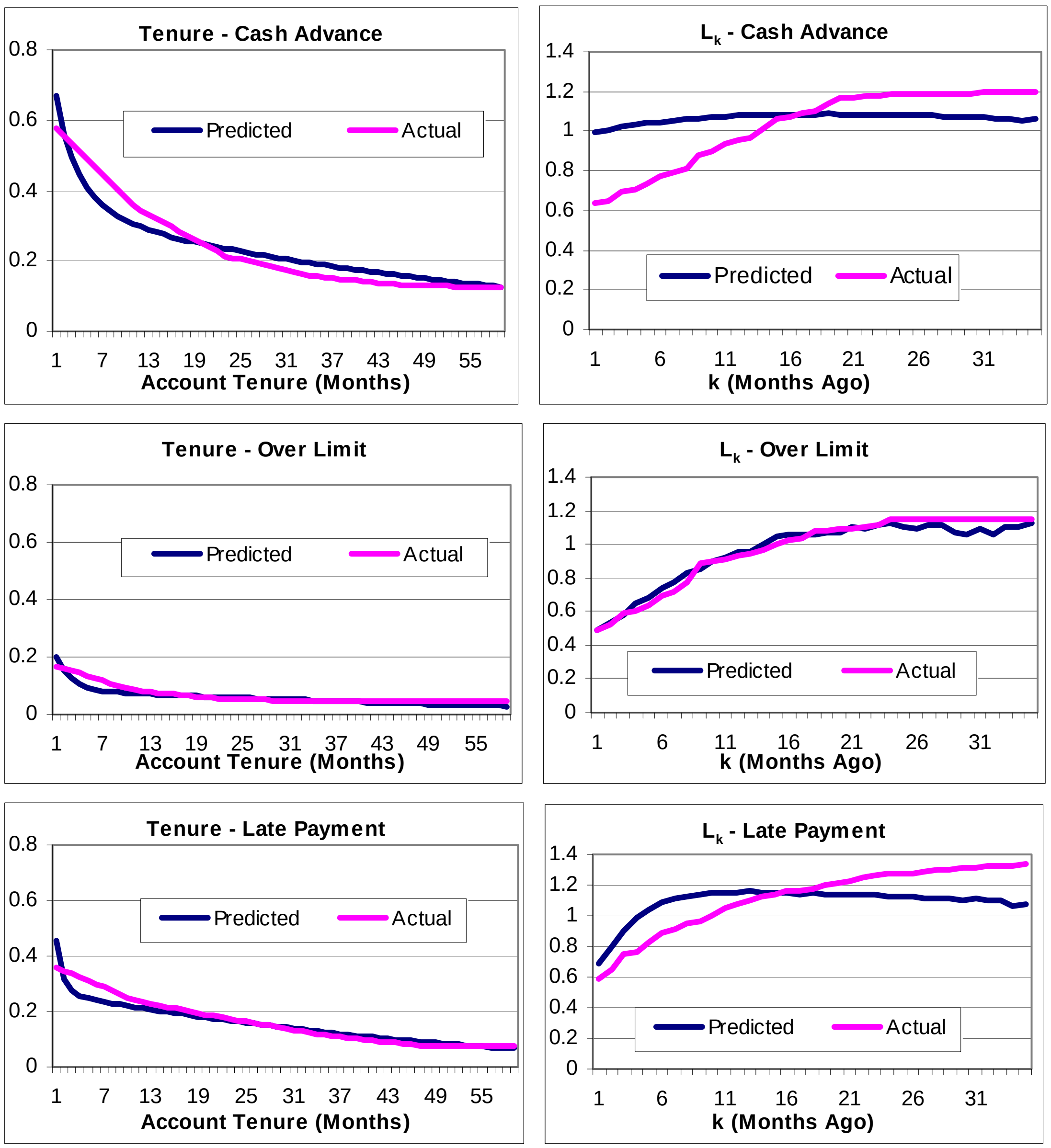


Figure 5: Demeaned and Normalized FICO Score, Behavior Score, Purchases, and Utilization Rates