LE
7/27/06
9:00 AM

Two Steps Forward, One Step Back:
The Dynamics of Learning and Backsliding

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

July 18, 2006

Abstract

We study learning with a three-year panel dataset representing four million monthly credit card statements. We focus on fees, since some observers claim that fees are not fully understood when consumers first open an account. In our data, new accounts generate effective monthly fee payments of $16 per month. Fee payments are negatively correlated at both low and high frequencies. Through negative feedback – i.e. paying a fee – consumers learn to avoid triggering fees in the future. These learning dynamics imply that monthly fee payments fall by 75% during the first four years of account life (controlling for account fixed effects). However, we find that consumers' hard-earned knowledge does not persist. We estimate that knowledge depreciates approximately 15% per month. Fee payments generate maximal learning when the fee payment was made last month. As negative feedback recedes into the distant past, consumers fail to notice it and their behavior tends to backslide. Like rational agents, consumers learn, but like myopic agents, consumers respond to recent events more than events that occurred just a few months ago.

JEL classification: D1, D4, D8, G2.

Keywords: forgetting, backsliding learning, shrouding, credit cards, fees, attention.

* Keith Ericson and Tom Mason provided outstanding research assistance. Laibson acknowledges financial support from the National Institute on Aging (R01-AG-1665). Sumit Agarwal: Federal Reserve Bank of Chicago. John Driscoll: Federal Reserve Board. Xavier Gabaix: Princeton University and NBER. David Laibson: Harvard University and NBER. 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. The authors are grateful to Murray Carbonneau, Stefano Dellavigna, Joanne Maselli, Benjamin Bolitzer, Peter Fishman, and seminar participants at the ASSA meetings, the Washington Area Finance Conference, and the Institute for Fiscal Studies conference on Consumer Behaviour and Bounded Rationality for helpful discussions and suggestions. All remaining errors are ours.

1 Introduction

Even if people can't solve a problem on the first attempt, most people will eventually learn to make the right choice. One may believe that over time, learning causes economies to converge to the equilibrium path that would arise if everyone were perfectly rational (Lucas 1986).

Many economists have shown that learning occurs in the laboratory (Camerer 2003). In such a controlled environment it is easy to measure learning since the experimenter can precisely observe behavior and can directly control incentives and feedback. It is also important to understand how learning arises in the field. In this paper, we study learning with a panel dataset that contains three years of credit card records, representing 120,000 consumers. For these accounts, we know what financial incentives the account holders face. We can track account holders' purchases, payments, and fees. And we know what feedback the account holders receive on their monthly credit card statements.

We focus our analysis on credit card fees — i.e. late payment fees, over limit fees, and cash advance fees — since academic and media observers have argued that new customers do not anticipate how frequently they will end up paying these fees.[1] We want to know whether credit card holders get more sophisticated with experience, effectively learning enough about fees to avoid triggering them.

Fee payments *are* large immediately after consumers open an account. New accounts generate direct monthly fee payments that average \$16 per month. 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 that may increase debt service costs by hundreds of dollars per year.[2]

As new consumers receive negative feedback — i.e., as they pay fees and experience resulting increases in their rate of interest — they learn to avoid triggering fees in the future. Controlling for person fixed effects, fee payments fall by 75 percent during the first four years of account life. We show that this learning is driven by feedback. Making a late payment (and consequently paying

[1] For example, Frontline reports that "The new billions in revenue reflect an age-old habit of human behavior: Most people never anticipate they will pay late, so they do not shop around for better late fees." See Gabaix and Laibson (2006) for a theoretical model of such effects. (http://www.pbs.org/wgbh/pages/frontline/shows/credit/more/rise.html)

[2] 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 prompt a report to the credit bureau, adversely affecting the card holder's credit accessability and creditworthiness. The average consumer has 4.8 cards and 2.7 actively used cards. Adverse credit reports may raise the interest rates on all of these cards.

a fee), sharply reduces the probability of another late payment in the subsequent month. As a consequence, fee payments are negatively correlated over time (at all frequencies).

This effect may be driven though many different channels. Consumers may learn that fees exist when they are forced to pay them. Alternatively, consumers may simply pay more attention to their credit card account when they have recently paid fees. Whatever the mechanism, paying substantial fees is not a feature of steady-state behavior or price discrimination among rational consumers.

We also find that the learning mechanism is not monotonic. Card holders act as if their hard-earned knowledge depreciates. A late payment charge from last month is more influential than an identical charge that was paid a year ago. Formally, the monthly hazard rate of a fee payment *increases* as previous fee payments recede further into the past (holding all else equal). Using a standard depreciation model, we estimate a knowledge depreciation rate of approximately 15 percent per month.

Our credit card data implies that card holder attention waxes and wanes as a function of the recency of negative feedback. Recent fee payments lead consumers to be highly vigilant, depressing fee payments in the immediate future. But a long string of success leads consumers to relax their vigilance and grow inattentive. This leads to new fee payments and another round in the cycle of learning. This dynamic process produces a stochastic steady state in which credit card payments are negatively correlated across time. Credit card holders need occasional failures to remind them to stay on the straight and narrow path.

Our findings imply that learning is very powerful, but that backsliding partially offsets learning, producing an important source of variability in consumer behavior. Other examples of such dynamics are easy to imagine. After visiting a doctor's office, a patient may vigilantly take her heart medicine. But a month later, the the doctor's warnings will lose their force and the patient's compliance will deteriorate. Only another pep talk will (temporarily) resurrect the patient's commitment to her medical regimen.

Many important behaviors appear to fluctuate in this way (at least in our own personal experience). After witnessing an accident, we temporarily driver more carefully. After visiting a relative with diabetes, we temporarily eat healthier meals. We respond to threats that are psychologically salient. Recent events loom largest in our minds, playing a disproportionate role in influencing our behavior.

We conclude that decision-makers do learn, but they learn best when the feedback was recent.

The fact that recent information matters more than older information tends to make behavior volatile. As yesterday's lessons recede into the past, we cease to attend to them and our behavior tends to backslide.

The paper has the following organization. Section 2 presents a simple model of learning and forgetting. Section 3 summarizes our data and presents our basic empirical results. Section 4 discusses an extension of the model and presents related empirical evidence. Section 5 presents a literature review. Section 6 concludes.

2 A simple model of learning and backsliding

In this section we describe a simple learning model. This model includes three components: a stock of feedback, a dynamic updating equation, and a mapping from 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, evolves according to:

$$F_t = f_t + (1 - \delta) F_{t-1}. \tag{1}$$

In this dynamic updating equation, $\delta \in [0, 1]$ represents the depreciation rate of the stock of feedback. We refer to this as "depreciation" but δ captures many related effects including recency bias, salience, forgetting, or any other form of temporal backsliding.[3]

If $\delta = 0$, there is no depreciation, and if $\delta = 1$, there is full depreciation after one period. Solving for F_t,

$$F_t = \sum_{i=0}^{t} (1 - \delta)^i f_{t-i}. \tag{2}$$

We assume that past fee payments drive down future fee payments (through a learning mechanism, like reinforcement).[4] Controlling for person fixed effects, the more fees you have paid in the past,

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

[4] E.g., see Camerer (2003) and Sutton and Barto (1998).

the more likely you are to avoid paying fees in the future:

$$E_{t-1}[f_t] = a - \beta \frac{F_{t-1}}{1+\gamma F_{t-1}}. \tag{3}$$

with $\gamma \geq 0$. If $\gamma > 0$, then $\lim_{F\to\infty} \frac{F}{1+\gamma F} = 1/\gamma$, and learning saturates as F gets large. Specifically, additional fee payments do not generate any more fee avoidance for large F. Whatever the value of γ, the first-order effect of a change in F_{t-1} on $E_{t-1}[f_t]$ is $-\beta$ (when the Taylor expansion is taken around $F_{t-1} = 0$). Parameter β captures the strength of learning. When β is large, past feedback reduces the expected current rate of fee payment.

To microfound equation (3), consider the following simple memory model. The baseline probability of paying a fee is a. In some states, the memory system reminds the agent about fees (enabling the agent to avoid paying the fee). This reminder occurs with probability

$$\frac{\beta' F_{t-1}}{1+\gamma F_{t-1}}, \tag{4}$$

with $\beta' \leq \gamma$. So the chance of paying a fee is

$$a\left(1 - \frac{\beta' F_{t-1}}{1+\gamma F_{t-1}}\right).$$

Hence, the probability of paying a fee is $a - \beta\frac{F_{t-1}}{1+\gamma F_{t-1}}$ if $a\beta' = \beta$.

We view Eq. (4) as a representation of the psychology of memory. This equation implies that events that have happened relatively frequently are easier to remember.[5]

Empirically, we estimate that γ is close to 0, so it is useful to point out the implications for this tractable limiting case. The mean expected values, $\overline{F}_t = E_0[F_t]$ and $\overline{f}_t = E_0[f_t]$, satisfy:

$$\overline{F}_t = (1-\delta)\,\overline{F}_{t-1} + \overline{f}_t \tag{5}$$

$$\overline{f}_t = a - \beta\overline{F}_{t-1}. \tag{6}$$

Hence,

$$\overline{F}_t = (1-\delta-\beta)\,\overline{F}_{t-1} + a. \tag{7}$$

[5] The same microfoundation applies to (13). Each memory system can remind the agent of the fees.

The steady state values $\overline{F}_\infty = \lim_{t\to\infty} \overline{F}_t$ and $\overline{f}_\infty = \lim_{t\to\infty} \overline{f}_t$ satisfy $\overline{F}_\infty = (1-\delta-\beta)\,\overline{F}_\infty + a$, so

$$\overline{F}_\infty = \frac{a}{\beta+\delta}$$
$$\overline{f}_\infty = \frac{\delta a}{\beta+\delta}.$$

Since, $\overline{F}_t - \overline{F}_\infty = (1-\delta-\beta)\left(\overline{F}_{t-1} - \overline{F}_\infty\right)$, we can write,

$$\overline{F}_t = \overline{F}_\infty - (1-\delta-\beta)^t \left(\overline{F}_\infty - \overline{F}_0\right) \tag{8}$$
$$\overline{f}_t = \overline{f}_\infty - (1-\delta-\beta)^t \left(\overline{f}_\infty - \overline{f}_0\right). \tag{9}$$

Hence, the average frequency of events declines exponentially from its initial value $\overline{f}_t = a$ to its steady state value $\overline{f}_\infty = \frac{\delta a}{\beta+\delta}$. We estimate the strength of learning (β), and the rate of depreciation (δ). A high rate of learning reduces the steady state frequency of fee payments, since $\overline{f}_\infty = \frac{\delta a}{\beta+\delta}$. By contrast, a high value of forgetting increases the steady state value of $\overline{f}_\infty$.

3 Results

Section 2 describes a learning model in which more past fee payments lead to fewer current fee payments. In this section, we estimate this model, controlling for other phenomena that might lead current fee payments to be negatively correlated with past fee payments.

3.1 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) 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.[6]

1. **Late Fee**: A late fee of between $30 and $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 'penalty pricing' by raising the APR to over 24 percent. The bank may also choose to report late payment to credit bureaus, adversely affecting consumers' FICO scores. 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 promotional) level.

2. **Over Limit Fee**: An over limit fee, also of between $30 and $35, is assessed the first time the borrower exceeds his or her credit limit. The same penalty pricing as in the late fee is imposed.

3. **Cash Advance Fee**: A cash advance fee of the greater of 3 percent of the amount advanced, or $5, is levied for each cash advance on the credit card. Unlike the first two fees, this fee can be assessed many times per month. It does not cause the imposition of penalty pricing on purchases or debt. However, the APR on cash advances is typically greater than that on purchases, and is usually 16 percent or more.

3.2 Data Summary

Before we turn to an estimation of the model discussed above, we first present a reduced form analysis of the data.

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

[6] 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 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.

account effects (like the account holder's FICO score). The data plotted in Figure 1 is generated by estimating,

$$
\begin{aligned}
(10) \qquad f_{i,t}^{j} \quad = \quad & \alpha + \phi_i + \psi_{time} + Spline(Tenure_{i,t}) \\
& + \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} \\
& + \gamma_1 FICO_{i,t-3} + \gamma_2 Behave_{i,t-3} + \gamma_3 Util_{i,t-1} + \epsilon_{i,t}.
\end{aligned}
$$

$f_{i,t}^{j}$ is a dummy variable which 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 — $f_{i,t}^{Late}$ — over limit fees — $f_{i,t}^{Over}$ — and cash advance fees — $f_{i,t}^{Advance}$. 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[7] 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; $FICO_{i,t-3}$ measures quarterly[8] credit risk (FICO, for 'Fair, Isaac and Company'); $Behave_{i,t-3}$ is a quarterly proprietary credit score based on patterns of payment and debt, designed to predict aspects of account payment behavior not captured by $FICO$; $Util_{i,t}$, 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).[9] 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 50% of all accounts to 11% of all accounts. The frequency of late fee payments drops from 32% to 9%. Finally, the frequency of over limit fee payments drops from 14% to 3%.

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

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

[8] $FICO$ and $Behave$ are lagged three periods because they are only collected quarterly.

[9] Tenure in all figures starts at month four since we lag both the (quarterly) FICO and Behavior scores in the regressions by three months, and such scores are not available before account opening.

plotted in Figure 2 is generated by estimating,

$$\begin{aligned}V_{i,,t}^{j} &= \alpha+\phi_i+\psi_{time}+Spline(Tenure_{i,t})\\&\quad+\eta_1 Purchase_{i,t}+\eta_2 Active_{i,t}+\eta_3 BillExist_{i,t-1}\\&\quad+\gamma_1 FICO_{i,t-3}+\gamma_2 Behave_{i,t-3}+\gamma_3 Util_{i,t-1}+\epsilon_{i,t}.\end{aligned}$$

$V_{i,t}^{j}$ 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 \$6.29 per month in cash advance fees, \$5.29 per month in late fees, and \$2.25 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. As in Figure 1, Figure 2 shows that the average value of fee payments declines rapidly with account tenure.

Finally, note that both Figures 1 and 2 begin plotting data at 4 months of tenure (due to the presence of lagged right-hand-side variables). Hence, the effect of learning is understated in these figures, since the months with the least tenure are truncated.

Figure 3 reports the expected frequency of a fee payment as a function of the timing of lagged fee payments. For Figure 3, we estimate:

$$\begin{aligned}f_{i,t}^{j} &= \alpha+\phi_i+\psi_{time}+\sum_{s=1}^{\infty}\left[f_{i,t-s}^{j}\cdot Spline(s)\right]\\&\quad+\eta_1 Purchase_{i,t}+\eta_2 Active_{i,t}+\eta_3 BillExist_{i,t-1}\\&\quad+\gamma_1 FICO_{i,t-3}+\gamma_2 Behave_{i,t-3}+\gamma_3 Util_{i,t-1}+\gamma_4 Tenure_{i,t}+\epsilon_{i,t}.\end{aligned}$$

$Spline(s)$ is a spline[10] that takes lag s as its argument. $Tenure_{i,t}$ is account tenure. All other variables are defined above.

Figure 3 plots $E\left[f_{i,t}^{j}\right]$ for different fee paying histories. Each point on the horizontal axis corresponds to a history in which a fee was paid on that lag (and no other lags). So movement along the horizontal axis corresponds to a change in the timing of lagged fee payments (holding the total number of fee payments fixed). In all three fee categories, the expected frequency of fee payments is substantially lower — by about 10 percentage points — one month after having paid a fee than it is one year after paying a fee.

[10] The spline has knots at lags one, two, three, five, seven and nine.

Figure 4 reports the expected value of a fee payment as a function of the timing of lagged fee payments. For Figure 4, we estimate:

$$\begin{aligned} V_{i,,t}^{j} &= \alpha + \phi_i + \psi_{time} + \sum_{s=1}^{\infty} \left[f_{i,t-s}^{j} \cdot Spline(s) \right] \\ &+ \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} \\ &+ \gamma_1 FICO_{i,t-3} + \gamma_2 Behave_{i,t-3} + \gamma_3 Util_{i,t} + \gamma_4 Tenure_{i,t} + \epsilon_{i,t}. \end{aligned}$$

Figure 4 plots $E\left[V_{i,t}^{j}\right]$ for different fee paying histories. Each point on the horizontal axis corresponds to a history in which a fee was paid at that lag (and no other lags). Again, movement along the horizontal axis corresponds to a change in the timing of lagged fee payments (holding the total number of fee payments fixed). In all three fee categories, the expected value of fee payments is substantially lower — by about 20% — one month after having paid a fee than it is one year after paying a fee.

Figure 5 plots conditional correlations between fee payments separated by different periods of time. For Figure 5, we separately estimate

$$\begin{aligned} f_{i,t}^{j} &= \alpha + \phi_i + \psi_{time} + \theta f_{i,t-k}^{j} \\ &+ \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} \\ &+ \gamma_1 FICO_{i,t-3} + \gamma_2 Behave_{i,t-3} + \gamma_3 Util_{i,t-1} + \gamma_4 Tenure_{i,t} + \epsilon_{i,t}. \end{aligned}$$

for values of k equal to 1, 3, 6, 9, 12, 15, 18, 21, and 24, and plot θ for each regression. Note that the values of θ are all negative and quite large. Fee payments are thus (conditionally) negatively autocorrelated at all time horizons.

3.3 Estimating the learning model

Our main regressions[11] take the form:

$$\begin{aligned} f^j_{i,,t} &= \alpha + \phi_i + \psi_{time} + \beta \frac{F^j(\delta)_{i,t-1}}{1+\gamma F^j(\delta)_{i,t-1}} \\ &\quad + \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} \\ &\quad + \eta_4 FICO_{i,t-3} + \eta_5 Behave_{i,t-3} + \eta_6 Util_{i,t} + \epsilon_{i,t}, \end{aligned}$$

$F^j(\delta)_{i,t-1}$ is the "depreciated" number of times fees of type j were paid through tenure $t-1$ (equation 1). Table 1 also reports the mean and standard deviation for the $F^j(\delta)_{t-1}$ variables. These moments are calculated at the estimated value of the rate of forgetting, δ, for each type j of fee.The first three terms are a constant, and account- and time-fixed effects. The coefficient on $\frac{F^j(\delta)_{i,t-1}}{1+\gamma F^j(\delta)_{i,t-1}}$, β, parameterizes learning, while γ parameterizes the degree of satiation.

The main parameters of interest are β, which represents the rate of learning,[12] δ, which determines the rate of forgetting, and γ, which determines the saturation of learning. The other regressors control for phenomena that might confound our analysis. The account fixed effects control for different propensities to pay fees over different account holders, and the time fixed effects for phenomena, such as business cycles, that might lead all account holders to have a different payment propensity at a particular time. The FICO and Behavior scores are intended to account for the effects of financial distress. A short period of household distress, for example as the result of illness, might cause borrowers to pay late fees for short periods of time. The utilization variable accounts for the fact that some consumers do not carry a card balance (in industry terminology, they are 'transactors'), and some do carry a balance ('revolvers'). Consumers with higher debt may have a greater propensity for fee payment.

We first estimate a version of the model in which there is no forgetting, that is, restricting $\delta = 0$, for all three kinds of fees The results are given in Table 2. Each column reports results for each of the three different types of fees. All coefficient are estimated very precisely, as would be expected given the very large number of observations (3.9 million statements). Row two reports the learning coefficient β in equation 3.

The saturation parameter γ in Eq. 3 is economically close to zero (though statistically distinct

[11] We estimate this equation with nonlinear least squares due to the large number of observations (3.9 million account statements).

[12] More precisely, β measures the rate at which stocks of fee-paying experience affects future behavior.

from zero). These low estimates imply that learning saturation only happens for very large values of F (e.g., F larger then $1/\gamma \approx 20$).

The estimated values of the learning coefficient shows that having paid fees in the past substantially reduces the propensity to pay fees now. For small values of F, paying a late fee in the past reduces the probability of paying a late fee today by 11 percentage points. Paying an over limit or cash advance fee also reduces the probability of subsequently paying such a fee by 11 percentage points.

As expected, temporarily having a higher FICO and Behavior Score — that is, being a temporarily less risky customer — substantially reduces the likelihood of paying a fee. The effects are comparable in magnitude to those of past fee payment: a one-standard deviation increase in FICO score reduces the propensity to pay a fee by about as much as having made one additional fee payment in the past.

Higher utilization — as measured by debt divided by the credit limit — also leads to greater fee payment. However, the effects are not large, suggesting little behavioral difference between revolvers and transactors.

This specification restricts learning effects to *not* exhibit a recency bias (i.e. $\delta = 0$). Thus a fee paid three years ago is forced to have exactly the same impact on current fee payment as a fee paid last month. To test this restriction, we re-estimate the specification allowing for a recency bias. The effects of past fee payments are now allowed to decline at a rate δ per month, as implied by equation 1. The estimates for the three types of fees are presented in Table 3.

All variables are again highly statistically significant. The estimate of the β learning parameters — the coefficient on $\frac{F^j(\delta)_{i,t-1}}{1+\gamma F^j(\delta)_{i,t-1}}$ — more than doubles. This is expected, since forgetting effects offset these stronger learning effects. The estimated value for the learning parameter now implies that paying a late fee in the previous month reduces the current propensity of paying a late fee by 21 percentage points. The effect is is stronger for the over limit fee (23 percentage points) and weaker for the cash advance fee (17 percentage points).

The depreciation parameter is economically very important.[13] The estimated values imply a depreciation rate of 15 percent per month for the late fee, 17 percent per month for the over limit fee, and 14 percent for the cash advance fee. So a fee paid one year ago has about 1/10 the impact on current fee payment as a fee paid last month.

This large depreciation effect explains why the learning effects (β) is so large. Depreciation

[13] Note that these parameters are part of the definition of the $N^j(\delta)_{i,t-1}$ variables.

makes this large learning effect rapidly fade away as fee payments recede into the past. For example, holding all else equal, the forgetting model predicts that paying an extra fee 12 months ago reduces the current propensity to pay a fee by only 2 percentage points. Hence the modest average learning effect can be better understood as a large short-run learning impact and a weak long-run learning impact.

4 Generalizations

4.1 A model with a short term and long term memory

Psychological studies in learning have identified at least two types of knowledge: knowledge held in short-term memory and knowledge held in long-term memory. McGaugh (2000), for example, presents comparisons of memory strength for several different memory systems. To capture such different effects, we introduce two decay parameter, δ_S (**s**hort term), and δ_L (**l**ong term), with $\delta_S > \delta_L$. We call:

$$F_t^S = F_t(\delta_S) = \sum_{i=0}^{t} (1-\delta_S)^i f_{t-i} \tag{11}$$

$$F_t^L = F_t(\delta_L) = \sum_{i=0}^{t} (1-\delta_L)^i f_{t-i} \tag{12}$$

The associated behavioral model is

$$E_{t-1}[f_t] = a - \beta_S \frac{F_{t-1}^S}{1+\gamma_S F_{t-1}^S} - \beta_L \frac{F_{t-1}^L}{1+\gamma_L F_{t-1}^L}, \tag{13}$$

where F_t^S represents short-term knowledge and F_t^L represents long-term knowledge. Those two dimensions of learning are important in our data.

Table 4 reports results from re-running the regression reported in Table 3, but now allowing for short-term and long-term memory. We include two $\frac{F_{t-1}}{1+\gamma F_{t-1}}$ terms. We do not restrict the choice of parameters, allowing the nonlinear least squares algorithm to freely pick the long- and short-term learning and forgetting parameters.

The regression results show evidence for the existence of two types of memory; both sets of new parameters are statistically significant, and have economically important magnitudes. The long-term memory parameters imply that a fee payment, to the extent that it is not forgotten, reduces the propensity to subsequently pay a fee by between 3 and 4 percentage points per month. Memory of such fee payments is very persistent, as they depreciate at a rate of around 2 percent per month.

By contrast, the effects of short-term memory produce sharp short-run reductions in the propensity to pay fees. Having paid a fee last month reduces the propensity to pay a fee subsequently by about 26-28 percentage points per month. However, memory of these fee payments fades quickly; the estimated rate of depreciation lies between 18 and 21 percent per month.

4.2 Discussion of alternative explanations not based on learning

The patterns that we have observed could be 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 (one inverse[14] measure of financial distress) correlate with account tenure. We find no such economically significant relationship. 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 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 appear to meaningfully change with account tenure.

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. We find no such economically significant relationship. We predict *Purchase* 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. Figure 6 plots the estimated spline.

[14] A high FICO score implies that the individual is a reliable creditor.

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 Util. 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 can not explain our learning dynamics, which are over an order of magnitude larger.

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

$$\Pr(f_t = 1) = \nu_t. \tag{14}$$

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,

$$f_t = \theta f_{t-1} + \text{controls.} \tag{15}$$

If (14) holds, then the regression coefficient is $\theta = cov(\nu_t, \nu_{t-1})/var(f_{t-1})$.

We run this regression, including all of our usual control variables: time- and account-fixed effects, a tenure spline, *Purchase*, *Active*, *BillExist*, *Behavior*, *FICO* and *Util*.

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

Results for the three types of fees are given in Table 5. 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.[15]

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 absolute value: $corr\,(\nu_t, \upsilon_{t-1}) \leq \theta = -0.75$.[16]

We think that such a very strong negative autocorrelation of monthly needs is very unlikely.[17] 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 postive autocorrelation in time needs and fee payments over the two months, not a negative one. Third, the time varying needs would have to produce higher than average fee payment in one month followed by 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.[18] 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 (15) is most plausibly explained by a recency effect – consumers become temporarily vigilant about fee avoidance immediately after paying a fee.

[15] There is a potential small sample bias (Nickell 1981), to which we thank Peter Fishman 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 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.

[16] It is easy to see that under (14), $cov\,(f_t, f_{t-1}) = cov\,(\nu_t, \nu_{t-1})$, 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.

[17] 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.

[18] Guvenen (200x)

4.3 The dynamics of learning and forgetting by consumer age

Figure 7 reports the dynamics of learning and forgetting by consumer age. For Figure 7, we estimate:

$$
\begin{aligned}
f^j_{i,,t} \quad = \quad & \alpha + \phi_i + \psi_{time} + \sum_{s=1}^{\infty} \left[\beta \frac{F^j(\delta)_{i,t-1}}{1 + \gamma F^j(\delta)_{i,t-1}} \cdot Spline(Age) \right] \\
& + \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} \\
& + \eta_4 FICO_{i,t-3} + \eta_5 Behave_{i,t-3} + \eta_6 Util_{i,t} + \epsilon_{i,t},
\end{aligned}
$$

where $Spline(Age)$ is a spline[19] that takes consumer age as its argument. The rest of the variables are as defined in the sections above. Figure 7 shows that, younger consumers learn slower. More interestingly, the figure shows that older consumers both learn slower and forget much faster.The estimated values imply a depreciation rate of 12-17 percent per month for consumer with ages 30 - 50 for all three fee types, but a depreciation rate of 25-45 percent per month for consumer with ages 65 and older for all three fee types.

5 Literature Review

Our paper is related to several branches of the literature. A number of researchers have written about consumer credit card use. Our work most closely overlaps with that of Agarwal *et al.* (2005), who use another large random sample of credit card accounts to show that, on average, borrowers choose credit card contracts that minimize their total interest costs net of fees paid. About 40 percent of borrowers choose suboptimal contracts that result in their paying avoidable interest costs. While some borrowers incur hundreds of dollars of such costs, most of these borrowers subsequently switch to cost-minimizing contracts. The results of our paper complement those of Agarwal *et al.* (2005), since we find evidence of learning to avoid fees and interest costs given a particular card contract. Massoud, Saunders and Scholnick (2006) use a panel dataset of credit card terms offered by banks to show that credit card penalty fees depend positively on consumer default rates at the bank, are negatively correlated with card APR, and appear to be unaffected by bank market share or average income level in the state in which the bank is located.

[19] The spline has knots for consumer age at every 10 years from age 30 through age 70.

Several researchers have looked at the response of consumers to low, introductory credit card rates ('teaser' rates), and at the persistence of otherwise high interest rates. Ausubel (1999) uses a panel dataset to document adverse selection in the response of consumers to credit card solicitations. He finds evidence that consumers overreact to credit card teaser rates and argues that this may occur because they underestimate the chance that they will borrow when the teaser rates expires. Shui and Ausubel (2004) show that consumers prefer credit card contracts with low initial rates for a short period of time to ones with somewhat higher rates for a longer period of time, even when the latter is *ex post* more beneficial. Consumers also appear 'reluctant' to switch contracts. DellaVigna and Malmendier (2004) theorize that financial institutions set the terms of credit card contracts to reflect consumers' poor forecasting ability over their future consumption. Bertrand *et al.* (2005) find that randomized changes in the "psychological features" of consumer credit offers affect adoption rates as much as variations in the interest rate terms. Ausubel (1991) hypothesizes that consumers may be over-optimistic, repeatedly underestimating the probability that they will borrow, thus possibly explaining the stickiness of credit card interest rates. Calem and Mester (1995) use the 1989 Survey of Consumer Finances (SCF) to argue that information barriers create high switching costs for high-balance credit card customers, leading to persistence of credit card interest rates, and Calem, Gordy and Mester (2005) use the 1998 and 2001 SCFs to argue that such costs continue to be important. Kerr and Dunn (2002) use data from the 1998 SCF to argue that having large credit card balances raises consumers' propensity to search for lower credit card interest rates. Kerr (2004) use SCF data to argue that banks offer better lending terms to consumers who are also bank depositors, and about whom the bank would thus have more information.

Other authors have used credit card data to evaluate more general hypotheses about consumption. Agarwal, Liu and Souleles (2004) use credit card data to examine the response of consumers to the 2001 tax rebates. Gross and Souleles (2002a) use credit card data to argue that default rates rose in the mid-1990s due to declining default costs, rather than a deterioration in the creditworthiness of borrowers. Gross and Souleles (2002b) find that increases in credit limits and declines in interest rates lead to large increases in consumer debt. Ravina (2005) estimates consumption Euler equations for credit card holders and finds evidence for habit persistence.

Our work is also related to research on learning. There are a number of papers that model individual decision making when there is bounded memory. Lehrer (1988) looks at asymptotic behavior and payoffs in repeated games when there is bounded recall. Aumann and Sorin (1989) show that bounded recall can result in near-optimal equilibria in games of common interest. Sarin

(2000) finds that, for a large class of decision rules, agents converge to playing maximin strategies. Piccione and Rubinstein (1997a, 1997b) and Aumann, Hart and Perry (1997a, 1997b) debate the utility of using absentmindedness as a modeling strategy.

There is a large body of results on learning in laboratory settings. A significant part of the work by economists focuses on players in a game learning about others' preferences and strategies; such results are less relevant in our setting, in which one player, the bank, is passive, and the other, the borrower, is learning about the rules of the game. Our setting does fit reinforcement models, in which previous payoffs or penalties affect current choices. McAllister (1991) find that reinforcement learning explains outcomes in some coordination games. Roth and Erev (1995) find that reinforcement models have difficulty explaining the speed of learning in ultimatum games. VanHuyck, Battalio and Rankin (2001) also find that reinforcement learning correctly predicts the direction of learning, but fails to predict the speed of learning in multiple-player games. Van Huyck, Battalio and Beil (1990, 1991) present experimental evidence for adaptive learning in coordination games; Crawford (1995) finds evidence for adaptive learning in their results. These and other related papers are surveyed in Camerer (2003). Some recent discussion in the psychology literature may be found in Anderson (2000) and Wixted (2004a, 2004b).

Our paper is more closely related to work on learning in field settings. Learning about the economic environment has been the focus of some recent macroeconomic work. Sargent (1999) uses least-squares learning models to explain Federal Reserve attempts to identify the structure of the Phillips curve, and resulting policy actions, during the increase subsequent decline in inflation during the 1970s and 1980s. Marimon and Sunder (1994) and Evans, Honkaphoja and Marimon (2001) study learning about monetary policy changes in overlapping generations models. Evans and Honkaphoja (2001) provide a survey of related work.

There has been much less research about learning in the field than in learning in laboratory settings, and little of this work has been about learning about rules or inherent characteristics of the economic environment. Ho and Chong (2003) estimate a model of consumer learning about product attributes, and find that it has substantially greater predictive power, with fewer parameters, than other leading models used by retailers to forecast unit demand. Lemieux and MacLeod (2000) look at the aftermath of an increase in the generosity of the Canadian unemployment insurance (UI) system in 1971. They find that the propensity to subsequently collect UI increased with a first-time exposure to this new system via an unemployment spell. They attribute part of this increase to both individual learning, through direct exposure to UI, and in part to social learning through peer-group exposure to UI. Barber, Odean and Strahlevitz (2004) find that individual investors tend to repurchase stocks that they previously sold for a gain, or have lost value subsequent to a prior sale, and to purchase additional shares of stocks that lost value subsequent to initial purchase. They attribute the first pattern to learning to actions that previously led to pleasure. McAfee and McMillan (1996) look at the evolution of strategies over time of the FCC airwaves auction; their findings are largely about players' learning about other players' preferences and strategies. Miravete (2003) finds evidence of learning in consumers' propensity to switch telephone calling plans to minimize monthly bill payments even for very small differences in cost.

There is a large literature that estimates rates of 'learning-by-doing,' in which firms learn how to improve their production processes in the act of producing. Recent contributions include Bahk and Gort (1993), who find that such learning within new plants largely occurs within the first ten years; Thompson (2001) and Thornton and Thompson (2001), who look at the production of Liberty ships during World War II; Gruber (1992) and Nye (1996), who look at the semiconductor industry; Zimmerman (1982), who examines the construction of nuclear power plants; and Barrios and Strobl (2004), who find evidence for spillovers across Spanish manufacturing firms. 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.

6 Conclusion

Using a three-year panel dataset representing 120,000 credit card consumers, we find that consumers learn how to avoid fees. In our data, new accounts generate direct monthly fee payments of \$16 per month and indirect costs that are even higher. Fee payments are negatively autocorrelated at all frequencies. The data implies that negative feedback — i.e. paying fees — teaches consumers to avoid triggering fees in the future. Controlling for person fixed effects, monthly fee payments fall by 75% during the first four years of account life.

We also find that consumers' hard-earned knowledge does not persist. In our basic specification, we estimate that past knowledge depreciates at a rate of 15% per month. As yesterday's fee-paying lesson recedes into the past, consumers tend to backslide.

These results simultaneously buttress and critique the rational actor model. Like rational agents, consumers learn, but like myopic agents, consumers tend to respond to recent events far more than events that occurred just a few months ago. Naive new customers quickly learn the ropes, but they need to keep relearning yesterday's lessons.

References

Agarwal, Sumit, Souphala Chomsisengphet, Chunlin Liu and Nicholas S. Souleles (2005). "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 (2006). "Financial Mistakes Over the Lifecycle," Manuscript, Harvard University.

Agarwal, Sumit, Chunlin Liu and Nicholas S. Souleles (2004). "The Reaction of Consumer Spending and Debt to Tax Rebates– Evidence from Consumer Credit Data." Manuscript, The Wharton School, University of Pennsylvania.

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 edn. Wiley, New York.

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

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

Aumann, Robert, Sergiu Hart and Motty Perry (1997a). "The Forgetful Passenger." *Games and Economic Behavior*, 20(1), pp. 117-120.

Aumann, Robert and Sylvain Sorin (1989). "Cooperation and Bounded Recall." *Games and Economic Behavior*, 1(1), pp. 5-39.

Ausubel, Lawrence M. (1991). "The Failure of Competition in the Credit Card Market." *American Economic Review*, 81(1), pp. 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), pp. 561-583.

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

Barber, Brad M., Terrance Odean and Michal Strahilevitz (2004). “Once Burned, Twice Shy: Naive 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), pp. 175-203.

Bénabou, Roland and Jean Tirole. “Self-Confidence and Personal Motivation.” *Quarterly Journal of Economics*, 117: 871-913.

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

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. “What’s Psychology 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) , pp. 1327-1336.

Calem, Paul S., Michael B. Gordy and Loretta J. Mester. “Switching Costs and Adverse Selection in the Market for Credit Cards.” Manuscript, Federal Reserve Bank of Philadelphia.

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), pp. 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( ), pp. 1750-1762.

DellaVigna, Stefano and Ulrike Malmendier (2004). “ContractTheory and Self-Control: Theory and Evidence.” *Quarterly Journal of Economics*, 119(2), pp. 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), pp. 1-31.

Gabaix, Xavier and David I Laibson (2006). “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets.” *Quarterly Journal of Economics*, 121(2), 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, pp. 149-185.

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

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

Haken, Hermann (1983) *Synergetics* 3rd edn Springer-Verlag, Berlin.

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

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.

Lehrer, Ehud (1988). "Repeated Games with Stationary Bounded Recall Strategies." *Journal of Economic Theory*, 46(1), pp. 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), pp. 139-170.

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), pp. 159-175.

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

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

Malmendier, Ulrike, and Devin Shanthikumar, "Are Small Investors Naive about Incentives?" mimeo, Harvard University and Stanford University, 2005.

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

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

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

Nickell, Stephen (1981). "Biases in Dynamic Models with Fixed Effects." *Econometrica*, 49(6), pp. 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), pp. 383-394.

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

Piccione, Michele and Ariel Rubinstein (1997b). "The Absent-Minded Driver's Paradox: Synthesis and Response." *Games and Economic Behavior*, 20(1), pp. .

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), pp. 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. "Time Inconsistency in the Credit Card Market."Manuscript, University of Maryland.

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), pp. 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), pp. 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), pp. 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), pp. 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), pp. 975-2005.

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

——— (2004b). "The Psychology and Neuroscience of Forgetting." *Annual Review of Psychology*, 55( ), pp. 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), pp. 297-310.

Appendix A: Description of the data

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 total 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.

Appendix B: Analysis of the two system learning model

We analyze the model in the case $\gamma = 0$. Taking expected values, the dynamics is:

$$F_{t+1}^{L} = (1 - \delta_L)\, F_t^L + f_t$$

$$F_{t+1}^{S} = (1 - \delta_S)\, F_t^S + f_t$$

$$f_t = a - \beta_L F_t^L - \beta_S F_t^L$$

We use the notation: $X_t = \left(F_t^L, F_t^S\right)', \beta = (\beta_L, \beta_S)$. Hence:

$$X_{t+1} = MX_t + b \tag{16}$$

$$f_{t+1} = a - \beta' \cdot X_t \tag{17}$$

with

$$M = \begin{pmatrix} 1-\delta_L-\beta_L & -\beta_S \\ -\beta_L & 1-\delta_S-\beta_S \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} a \tag{18}$$

Hence:

$$X_t = a\left(1-M\right)^{-1}\left(1-M^t\right)\begin{pmatrix} 1 \\ 1 \end{pmatrix} + M^t X_0 \tag{19}$$

This makes calculating impulse responses easy. A surprise fee at time 0 increases X_0 by $(1,1)'$, and so create a probability of fees at time t equal to:$-\beta' M^t (1,1)'$. The total amount paid decreases by the $\sum_{t=0}^{\infty} \beta' M^t (1,1)'$, i.e. $\beta' (1-M)^{-1} (1,1)'$. Empirically, this number of -0.73 for late fees. This means that a late fee paid now, creates a learning, that induces consumers to pay 0.73 less fee in the future.

The above structure generates the dynamics announced in (**??**), where λ_S and λ_L are the eigenvalues of $I - M$, i.e. are the root of:

$$\lambda^2 - (\delta_S + \beta_S + \delta_L + \beta_L)\,\lambda + (\delta_S\delta_L + \delta_S\beta_L + \delta_L\beta_S) = 0 \tag{20}$$

The two roots are $\lambda_L < \lambda_S$, which we call, respectively, the long term root, or slow root, and the short term root, or fast root.

One can show the following Propositions.

Proposition 1 *Assume that β_L and β_S are positive. The roots of the integrated system, λ_L and λ_S, satisfy*

$$\lambda_L < \min\left(\delta_L + \beta_L, \delta_S + \beta_S\right) \leq \max\left(\delta_L + \beta_L, \delta_S + \beta_S\right) < \lambda_S \tag{21}$$

To interpret the Proposition, consider the case $\delta_L + \beta_L \leq \delta_S + \beta_S$. The slow root of the integrated system is lower than root of the separate slow system ($\lambda_L < \delta_L + \beta_L$), while the fast root of the integrated system is higher than in the separate system ($\delta_S + \beta_S < \lambda_S$).

Proof. Call $\Phi(\lambda)$ the polynomial in (20). $\Phi(\delta_L + \beta_L) = \Phi(\delta_S + \beta_S) = -\beta_L \beta_S < 0$, which implies that that $\delta_L + \beta_L$ and $\delta_S + \beta_S$ are in (λ_L, λ_S). ∎

Proposition 2 *Set $\varepsilon = (\delta_L + \beta_L) / (\delta_S + \beta_S)$. In the limit $\varepsilon \ll 1$, one has:*

$$\lambda_S = \beta_S + \lambda_S + \frac{\beta_S \beta_L}{\beta_S + \delta_S} + O\left(\varepsilon^2\right) = \beta_S + \lambda_S + O(\varepsilon)$$

$$\lambda_L = \delta_L + \beta_L - \frac{\beta_S \beta_L}{\beta_S + \delta_S} + O\left(\varepsilon^2\right)$$

Proof. Immediate, from (20). ∎

This Proposition means that the fast root of the integrated system, λ_S, is equal to the root of the separate fast system, to a leading order. The slow root is the root of the slow system, lowered by a non-trivial amount.[20]

[20] The analysis is close to that of hierarchical systems with a slow "master" system and a fast "slave" system (Haken 1983).

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.

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 bak 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: Regression with Just Learning

	$f_{i,t}^{Late}$	$f_{i,t}^{Over}$	$f_{i,t}^{Advance}$
Intercept	0.2787*	0.1682*	0.3445**
	(0.1380)	(0.0723)	(0.0774)
$\frac{F_{i,t-1}}{1+\gamma F_{i,t-1}}$	-0.1083**	-0.1074**	-0.1099**
	(0.0323)	(0.0384)	(0.0421)
γ	0.0770*	0.0437**	0.0694*
	(0.0357)	(0.0150)	(0.0284)
$Purchase/100_{i,t}$	0.0135	0.0127*	0.0171*
	(0.0086)	(0.0059)	(0.0076)
$Active_{i,t}$	0.0067	0.0020**	0.0096*
	(0.0049)	(0.0008)	(0.0048)
$BillExist_{i,t-1}$	0.0573*	0.0167	0.0974
	(0.0243)	(0.0273)	(0.0673)
$Behave_{i,t-3}$	-0.0024**	-0.0012*	-0.0023**
	(0.0008)	(0.0006)	(0.0006)
$FICO_{i,t-3}$	-0.0013**	-0.0005	-0.0009*
	(0.0004)	(0.0004)	(0.0004)
$Util_{i,t}$	0.0041*	0.0041**	0.0080*
	(0.0017)	(0.0014)	(0.0037)
Adjusted R-squared	0.0469	0.0490	0.0506
No. of Obs.	3.9 million	3.9 million	3.9 million

Notes: This table reports the results of estimating $f_{i,,t}^{j} = \alpha + \phi_i + \psi_{time} + \beta\frac{F^j(\delta)_{i,t-1}}{1+\gamma F^j(\delta)_{i,t-1}} + \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} + \eta_4 FICO_{i,t-3} + \eta_5 Behave_{i,t-3} + \eta_6 Util_{i,t} + \epsilon_{i,t}$, where the first three terms are a constant, and account- and time- fixed effects. δ is set to zero for this regression. The coefficient on $\frac{F^j(\delta)_{i,t-1}}{1+\gamma F^j(\delta)_{i,t-1}}$, β, parameterizes learning, while γ parameterizes the degree of satiation. Some super- and subscripts have been suppressed in the first column to aid legibility. In particular, the depreciated number of fee payments $F^j(\delta)_{i,t-1}$ varies across regression specifications to correspond to the type of fee used as the dependent variable. 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 level.

Table 3: Regression with Both Learning and Backsliding

	$f_{i,t}^{Late}$	$f_{i,t}^{Over}$	$f_{i,t}^{Advance}$
Intercept	0.2572*	0.1427**	0.3386**
	(0.1167)	(0.0437)	(0.0678)
$\frac{F(\delta)_{i,t-1}}{1+\gamma F(\delta)_{i,t-1}}$	-0.2078**	-0.2280**	-0.1699**
	(0.0407)	(0.0445)	(0.0404)
δ	0.1519**	0.1693**	0.1394*
	(0.0356)	(0.0350)	(0.0693)
γ	0.0654**	0.0434**	0.0630**
	(0.0245)	(0.0165)	(0.0206)
$Purchase/100_{i,t}$	0.0134	0.0126*	0.0173*
	(0.0086)	(0.0056)	(0.0074)
$Active_{i,t}$	0.0064	0.0024**	0.0095*
	(0.0038)	(0.0008)	(0.0040)
$BillExist_{i,t-1}$	0.0582*	0.0157	0.0982
	(0.0246)	(0.0283)	(0.0570)
$Behave_{i,t-3}$	-0.0014**	-0.0007**	-0.0016**
	(0.0005)	(0.0003)	(0.0006)
$FICO_{i,t-3}$	-0.0008**	-0.0004*	-0.0005**
	(0.0003)	(0.0002)	(0.0002)
$Util_{i,t}$	0.0039*	0.0047**	0.0061*
	(0.0016)	(0.0017)	(0.0028)
Adjusted R-squared	0.0604	0.0629	0.0694
No. of Obs.	3.9 million	3.9 million	3.9 million

Notes: This table reports the results of estimating $f_{i,,t}^{j} = \alpha + \phi_i + \psi_{time} + \beta \frac{F^j(\delta)_{i,t-1}}{1+\gamma F^j(\delta)_{i,t-1}} + \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} + \eta_4 FICO_{i,t-3} + \eta_5 Behave_{i,t-3} + \eta_6 Util_{i,t} + \epsilon_{i,t}$, where the first three terms are a constant, and account- and time- fixed effects. The coefficient on $\frac{F^j(\delta)_{i,t-1}}{1+\gamma F^j(\delta)_{i,t-1}}$, β, parameterizes learning, while γ parameterizes the degree of satiation. Some super- and subscripts have been suppressed in the first column to aid legibility. In particular, the depreciated number of fee payments $F^j(\delta)_{i,t-1}$ varies across regression specifications to correspond to the type of fee used as the dependent variable. 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 level.

Table 4: Regression with Two Types of Memory

	$f_{i,t}^{Late}$	$f_{i,t}^{Over}$	$f_{i,t}^{Advance}$
Intercept	0.2522*	0.1380**	0.3283**
	(0.1078)	(0.0439)	(0.0678)
$\frac{F_L(\delta)_{i,t-1}}{1+\gamma_L F_L(\delta)_{i,t-1}}$	-0.0357**	-0.0307*	-0.0301**
	(0.0106)	(0.0153)	(0.0099)
$\frac{F_S(\delta)_{i,t-1}}{1+\gamma_S F_S(\delta)_{i,t-1}}$	-0.2604**	-0.2799**	-0.2704**
	(0.0842)	(0.0767)	(0.0922)
δ_L	0.0227**	0.0209**	0.0239**
	(0.0090)	(0.0059)	(0.0094)
δ_S	0.1792*	0.2024*	0.2083*
	(0.0888)	(0.0901)	(0.0912)
γ_L	0.0940	0.0760	0.0892*
	(0.0487)	(0.0399)	(0.0444)
γ_S	0.0559*	0.0413**	0.0583*
	(0.0241)	(0.0176)	(0.0274)
$Purchase/100_{i,t}$	0.0139	0.0132*	0.0181*
	(0.0084)	(0.0063)	(0.0078)
$Active_{i,t}$	0.0065	0.0025**	0.0092
	(0.0040)	(0.0007)	(0.0059)
$BillExist_{i,t-1}$	0.0578*	0.0159	0.0932*
	(0.0243)	(0.0284)	(0.0383)
$Behave_{i,t-3}$	-0.0011*	-0.0007*	-0.0014*
	(0.0005)	(0.0003)	(0.0006)
$FICO_{i,t-3}$	-0.0008**	-0.0003**	-0.0007
	(0.0002)	(0.0001)	(0.0004)
$Util_{i,t}$	0.0033*	0.0040*	0.0049*
	(0.0016)	(0.0018)	(0.0024)
Adjusted R-squared	0.0610	0.0653	0.0701
No. of Obs.	3.9 million	3.9 million	3.9 million

Notes: This table reports nonlinear least squares estimates of

$$f_{i,t}^{Late} = \alpha + \phi_i + \psi_{time} - \beta_S \frac{F_S^{Late}(\delta_S)_{i,t-1}}{1+\gamma_S F_S^{Late}(\delta_S)_{i,t-1}} - \beta_L \frac{F_L^{Late}(\delta_L)_{i,t-1}}{1+\gamma_L F_L^{Late}(\delta_L)_{i,t-1}} + \eta_1 Purchase_{i,t} +$$

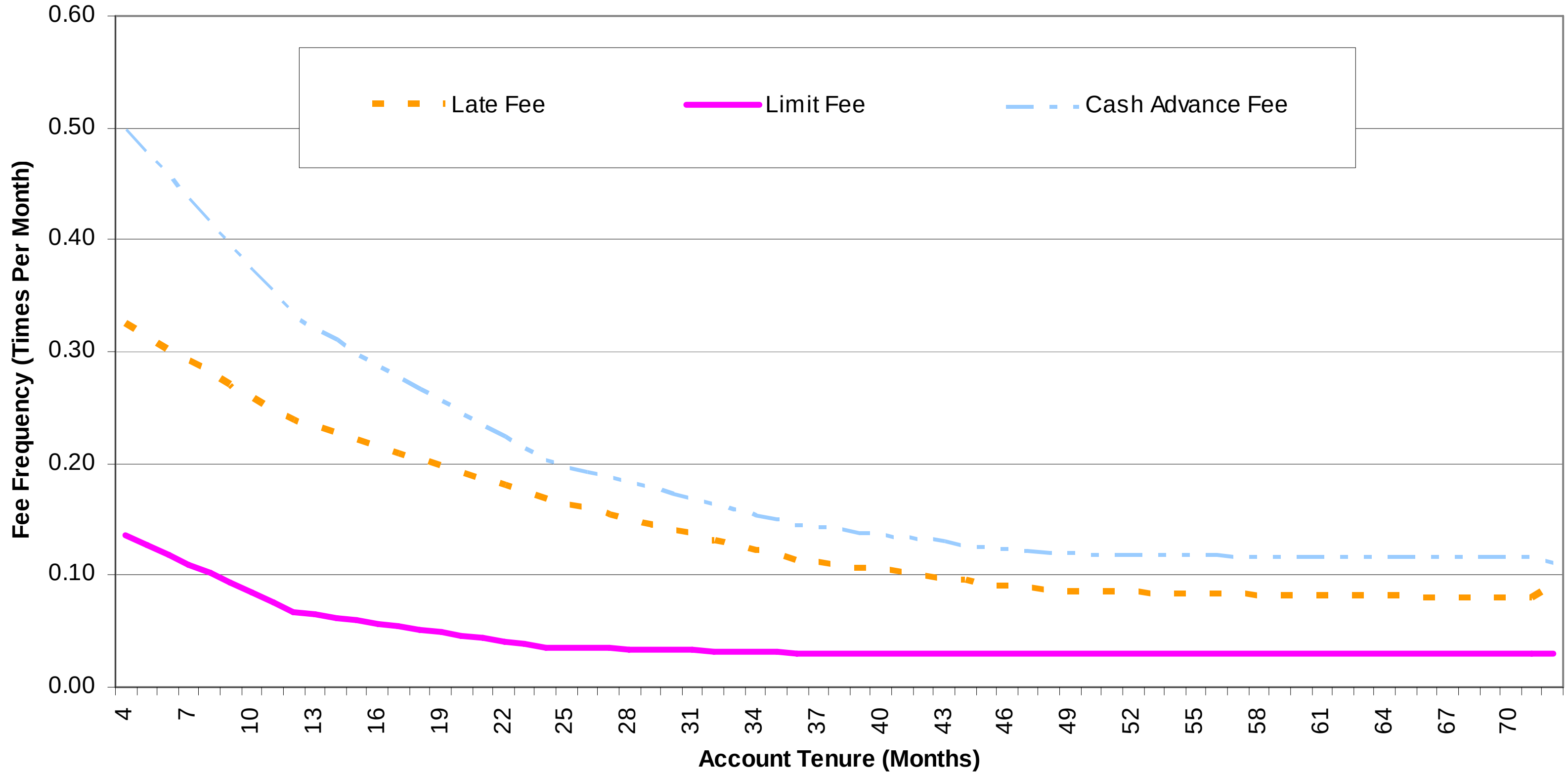
$\eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} + \eta_4 FICO_{i,t-3} + \eta_5 Behave_{i,t-3} + \eta_6 Util_{i,t} + \epsilon_{i,t}$, where the first three terms are a constant, and account- and time- fixed effects. S and L subscripts denote parameterizations of short- and long-term memory. Some super- and subscripts have been suppressed in the first column to aid legibility. In particular, the two types of depreciated number of fee payments $F^j(\delta)_{i,t-1}$ and the depreciation rates δ vary across regression specifications to correspond to the type of fee used as the dependent variable. 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 level.

Table 5: 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

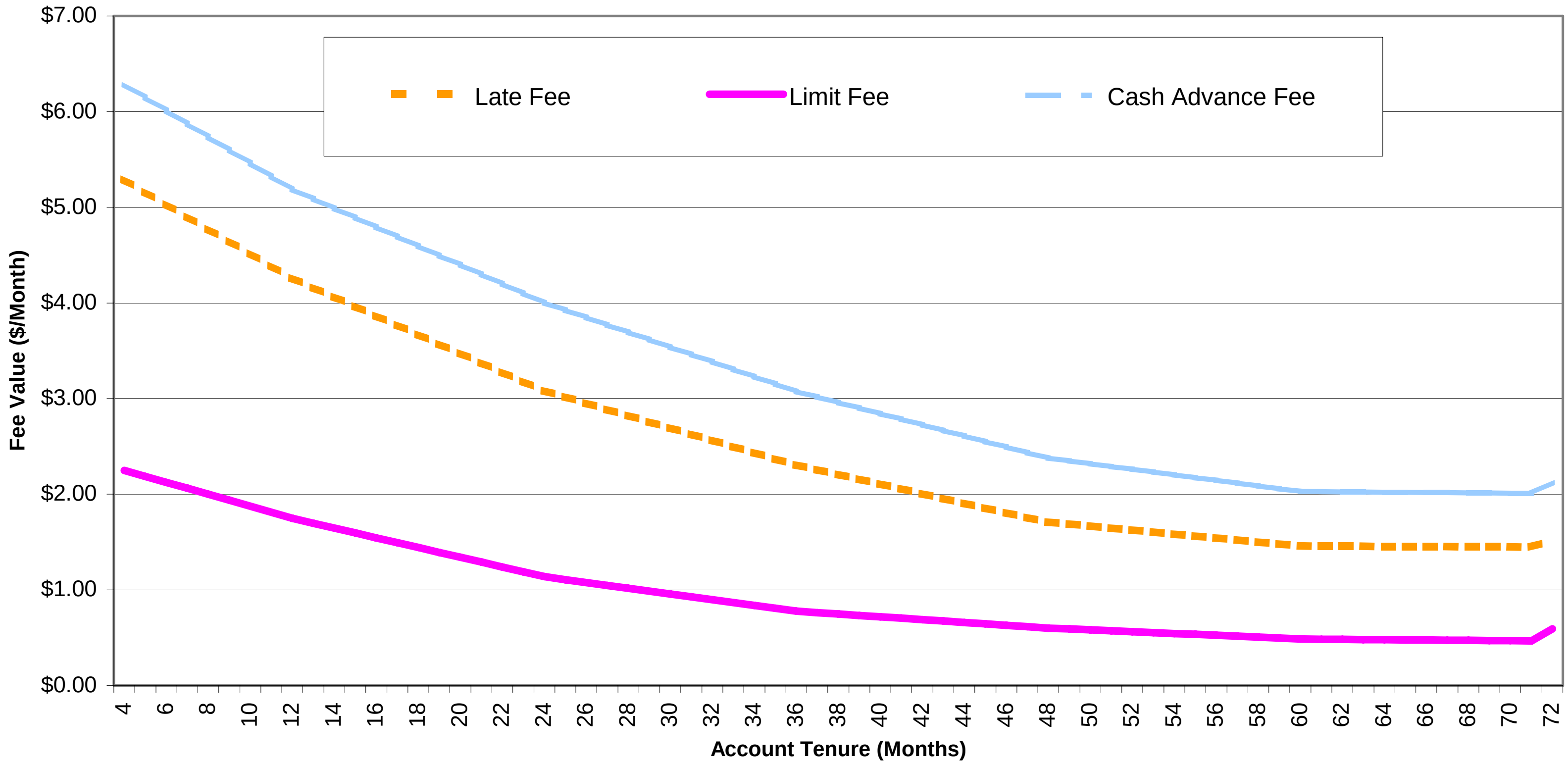
Notes: This table reports the results of estimating $f^j_{i,t} = \alpha + \phi_i + \psi_{time} + \theta f^j_{i,t-1} + Spline(Tenure_{i,t}) + \eta_1 Purchase_{i,t} + \eta_2 Active_{i,t} + \eta_3 BillExist_{i,t-1} + \eta_4 FICO_{i,t-3} + \eta_5 Behave_{i,t-3} + \eta_6 Util_{i,t} + \epsilon_{i,t}$, where the first three terms are a constant, and account- and time- fixed effects. Rows 3 through 8 report the coeficients 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 level.

Figure 1: Fee Frequency and Account Tenure



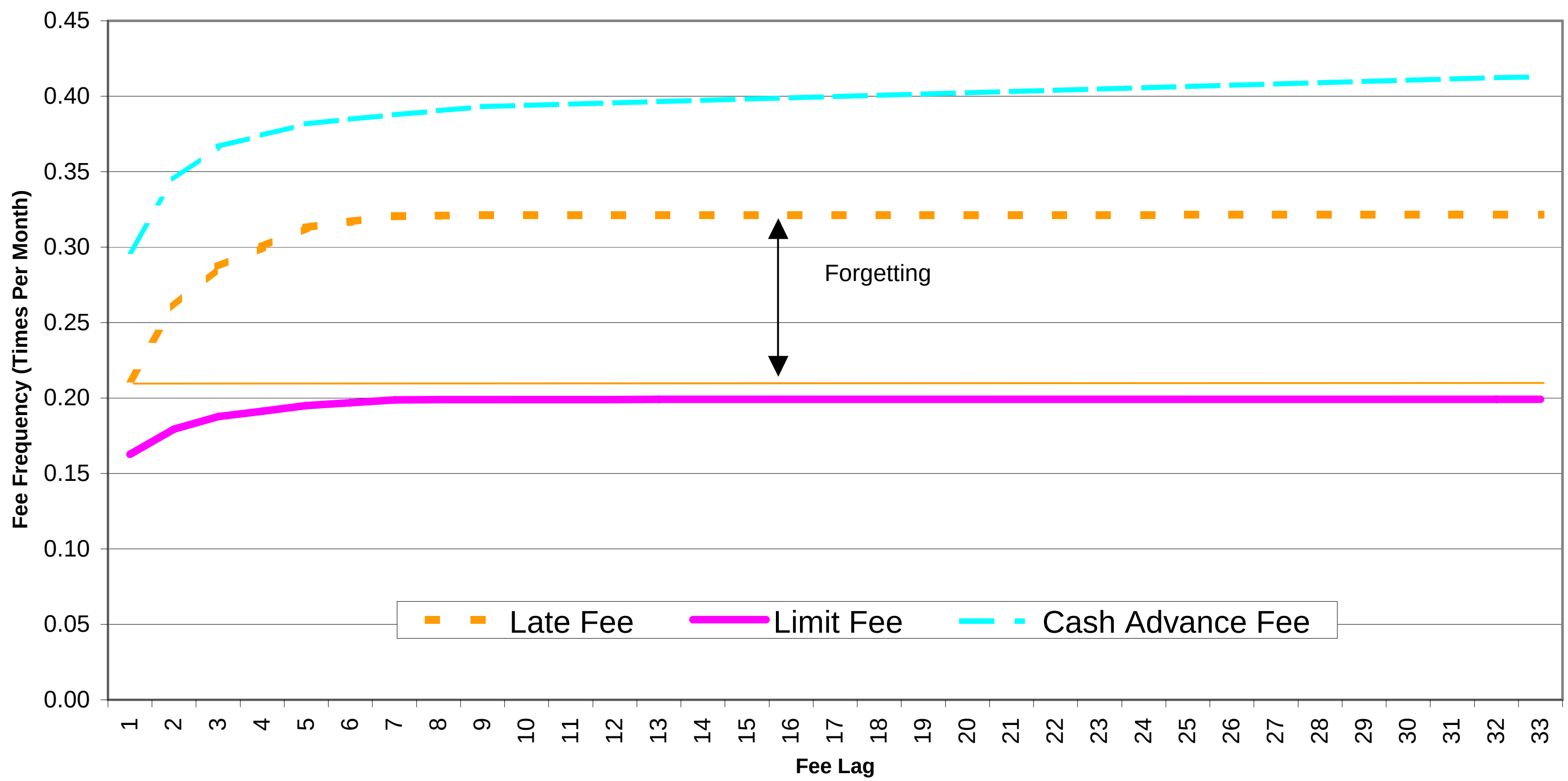
Notes: This figure plots the fitted values of a regression of fee frequency (times fee is paid per month) on a continuous piecewise linear function of account tenure (the function is a spline, with knots every twelve months, on the time since the account was opened), a constant, account- and time- fixed effects, and control variables (FICO score, Behavior score, utilization (Debt/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 fourth month because the FICO and Behavior scores, available only quarterly from account opening, are lagged three months in the regressions.

Figure 2: Fee Value and Account Tenure



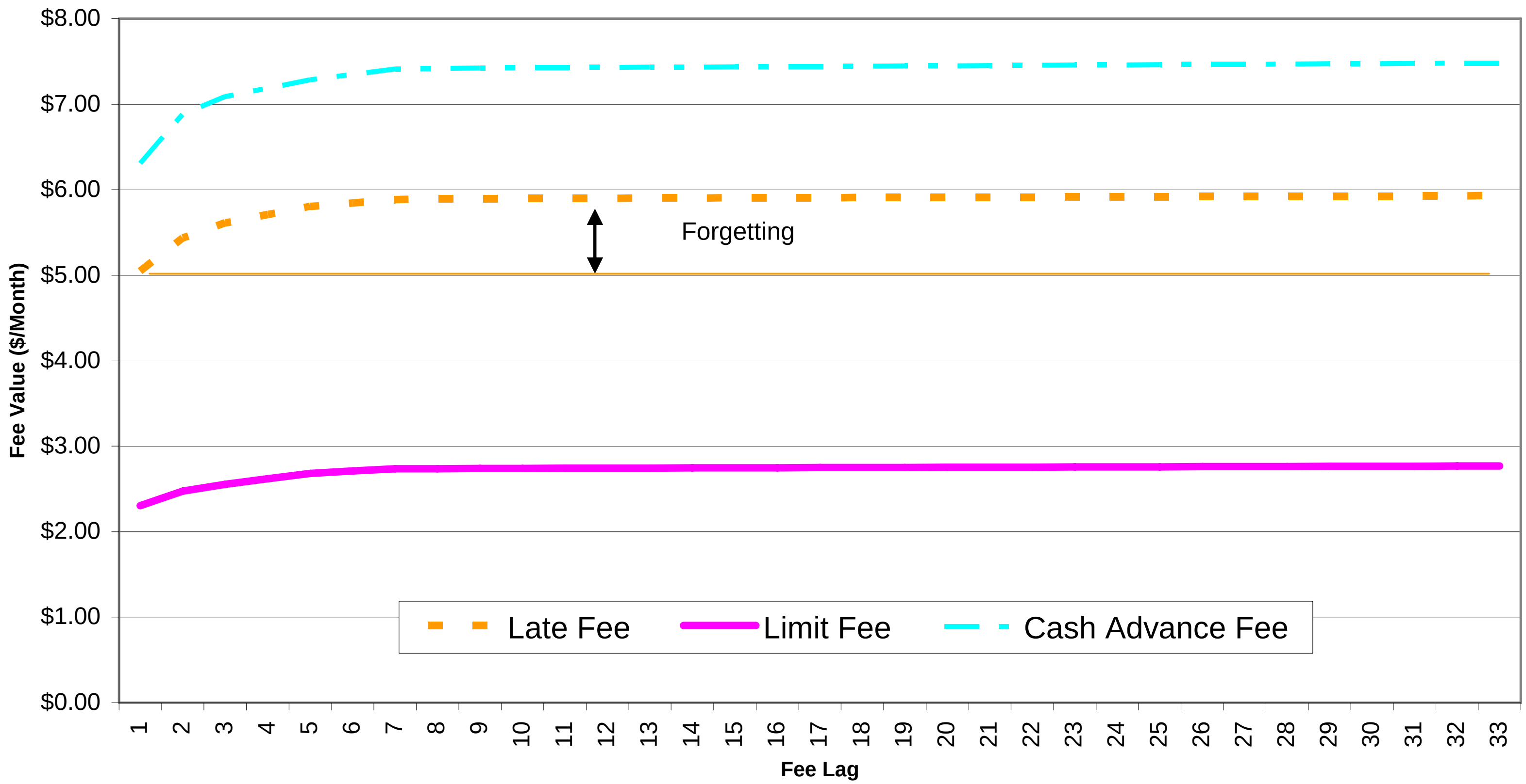
Notes: This figure plots the fitted values of a regression of fee value (dollars in fees paid per month) on a continuous piecewise linear function of account tenure (the function is a spline, with knots every twelve months, on the time since the account was opened), a constant, account- and time- fixed effects, and control variables (FICO score, Behavior score, utilization (Debt/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 fourth month because the FICO and Behavior scores, available only quarterly from account opening, are lagged three months in the regressions.

Figure 3: Fee Frequency and Time Since Last Fee Paid



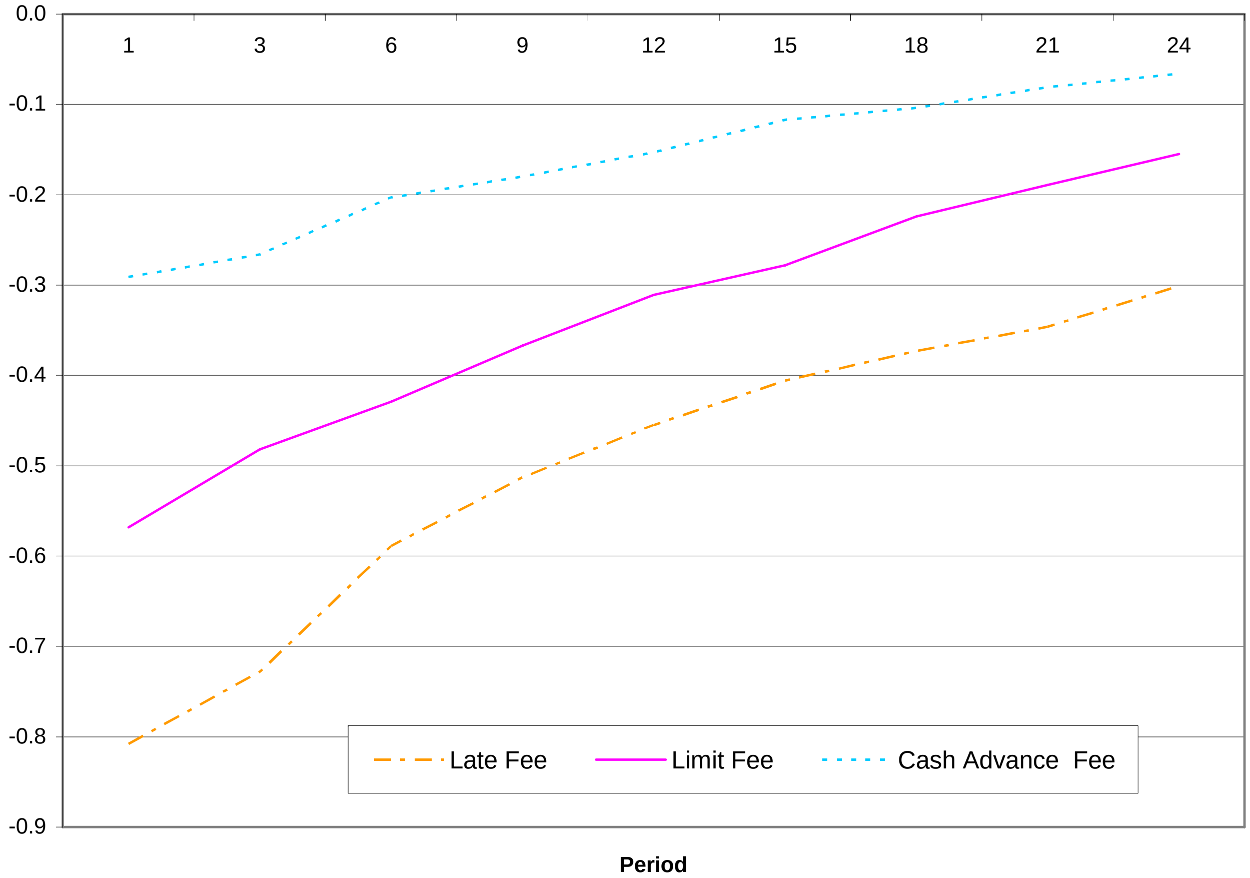
Notes: This figure plots the fitted values of a regression of fee frequency (times fee is paid per month) on a continuous piecewise linear function of time multiplied by a dummy for fee payment (the function is a spline, with knots at lags one, two, three, five, seven, and nine), a constant, account- and time- fixed effects, and control variables (FICO score, Behavior score, utilization (Debt/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 fourth month because the FICO and Behavior scores, available only quarterly from account opening, are lagged three months in the regressions.

Figure 4: Fee Value and Time Since Last Fee Paid



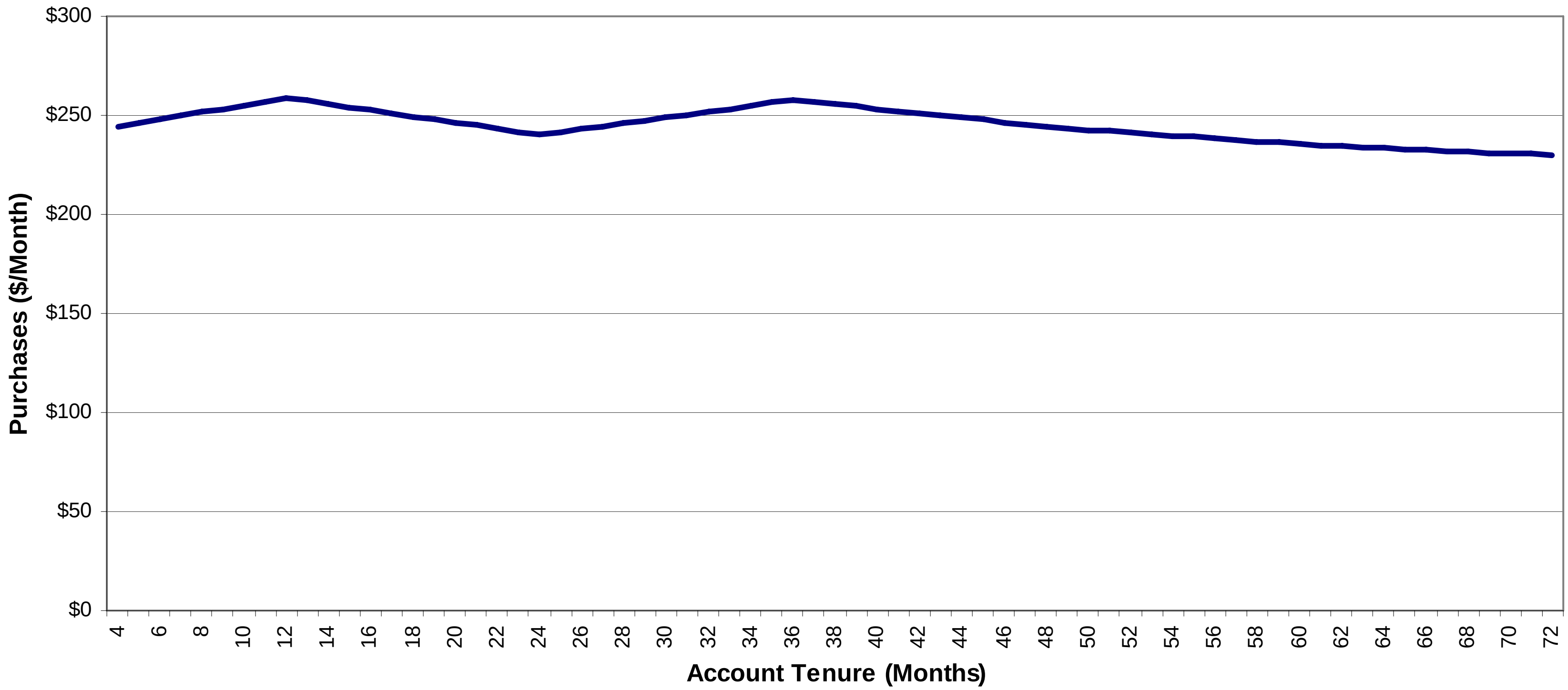
Notes: This figure plots the fitted values of a regression of fee value (dollars in fees paid per month) on a continuous piecewise linear function of time multiplied by a dummy for fee payment (the function is a spline, with knots at lags one, two, three, five, seven, and nine), a constant, account- and time- fixed effects, and control variables (FICO score, Behavior score, utilization (Debt/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 fourth month because the FICO and Behavior scores, available only quarterly from account opening, are lagged three months in the regressions.

Fig 5: Correlation of Fee Payments Between Periods t and t-k (k=1,3,6,9,12,15,18,21,24)



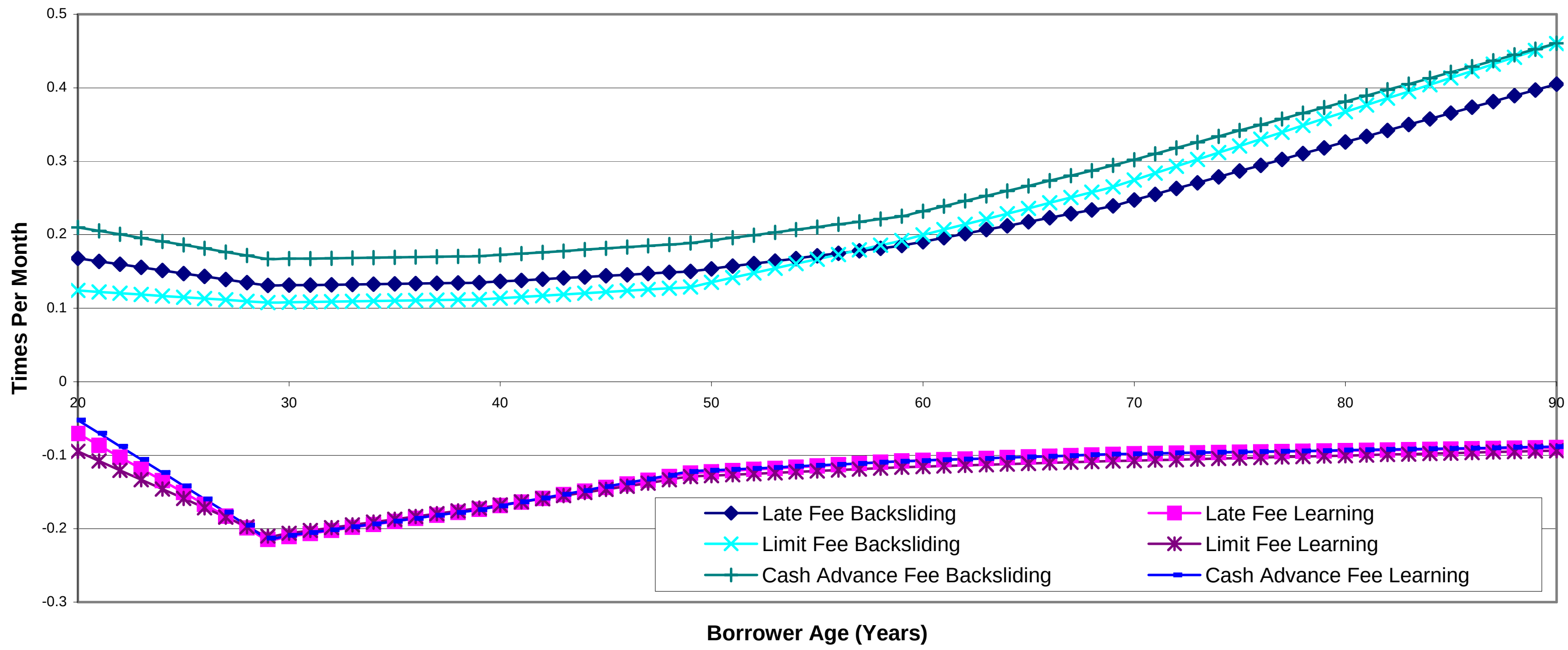
Notes: Each point on each curve represents the coefficient on lagged fee payment from a regression of current fee payment on fee payment k periods ago, a constant, account- and time- fixed effects, and the same control variables as in previous figures. k is allowed to take on values of 1,3,6,9,12,15,18,21 and 24.

Figure 6: Purchases and Account Tenure



Notes: This figure plots the fitted values of a regression of monthly purchases (in dollars) 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 (FICO score, Behavior Score and Utilization (Debt/Limit)). 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 fourth month because the FICO and Behavior scores, available only quarterly from account opening, are lagged three months in the regressions.

Figure 7: Rates of Learning and Backsliding by Age



Notes: This figure reports estimates of the rates of learning and backsliding in the regressions reported on Table 3, β and δ, where those coefficients have been allowed to vary in a piecewise linear way with age (the function is a spline, with knots at every ten years from ages 30 through 70). All regressions contain the same control variables as in Table 3. The top three, positive, coefficients plot rates of backsliding for the three different types of fees paid, while the bottom three, negative, coefficients plot rates of learning. Learning rates are negative because they measure the rate at which past fee payments reduce the current propensity to pay a fee. Each point on the learning rates graphs report the impact of a fee payment on the propensity to pay the same.