NBER WORKING PAPER SERIES

ESTIMATING DISCOUNT FUNCTIONS WITH CONSUMPTION CHOICES OVER THE LIFECYCLE

David Laibson
Andrea Repetto
Jeremy Tobacman

Working Paper 13314
http://www.nber.org/papers/w13314

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2007

We received helpful advice from Alberto Alesina, Orazio Attanasio, Felipe Balmaceda, Robert Barro, John Campbell, Christopher Carroll, Pierre-Olivier Gourinchas, Cristóbal Huneeus, Greg Mankiw, Jonathan Parker, Ariel Pakes, Raimundo Soto, Samuel Thompson, Nicholas Souleles, Motohiro Yogo, two referees and our editor Mark Gertler at the AER, and seminar audiences at the University of California-Berkeley, the Central Bank of Chile, Harvard University, NBER, the University of Minnesota, the University of Chile, Stanford University, and Princeton University. Laibson acknowledges financial support from the National Science Foundation (SBR-9510985), the National Institute on Aging (R01-AG-16605), and the Olin Foundation; Repetto, from DID-Universidad de Chile, FONDECYT (1990004), Fundación Andes and an institutional grant to CEA from the Hewlett Foundation; and Tobacman from a National Science Foundation Graduate Research Fellowship. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Estimating Discount Functions with Consumption Choices over the Lifecycle
David Laibson, Andrea Repetto, and Jeremy Tobacman
NBER Working Paper No. 13314
August 2007
JEL No. D91,E21

ABSTRACT

Intertemporal preferences are difficult to measure. We estimate time preferences using a structural buffer stock consumption model and the Method of Simulated Moments. The model includes stochastic labor income, liquidity constraints, child and adult dependents, liquid and illiquid assets, revolving credit, retirement, and discount functions that allow short-run and long-run discount rates to differ. Data on retirement wealth accumulation, credit card borrowing, and consumption-income comovement identify the model. Our benchmark estimates imply a 40% short-term annualized discount rate and a 4.3% long-term annualized discount rate. Almost all specifications reject the restriction to a constant discount rate. Our quantitative results are sensitive to assumptions about the return on illiquid assets and the coefficient of relative risk aversion. When we jointly estimate the coefficient of relative risk aversion and the discount function, the short-term discount rate is 15% and the long-term discount rate is 3.8%.

David Laibson
Department of Economics
Littauer M-14
Harvard University
Cambridge, MA 02138
and NBER
dlaibson@harvard.edu

Andrea Repetto
University of Chile
arepetto@dii.uchile.cl

Jeremy Tobacman
Department of Economics
Manor Road Building
University of Oxford
Oxford OX1 3UQ
jeremy.tobacman@economics.ox.ac.uk

1 Introduction

Though intertemporal preferences play a critical role in most important economic decisions, economists have not identified a reliable method for measuring them (Frederick, Loewenstein and O'Donoghue 2002). The vast majority of research on time preferences has used laboratory studies in which the experimenter controls the choices that subjects face. Laboratory experiments often ask subjects to weigh immediate rewards against delayed rewards. A typical study asks subjects if they would prefer $\$X$ now or $\$Y$ at a specified future date.

Despite the advantages of controlled laboratory experimentation, such studies may confound time preferences with other considerations, like the trustworthiness of the experimenter, the abstract nature of the laboratory task, or the outside investment options of the subject. It is not clear whether laboratory experiments measure the discount function, market interest rates, curvature of the utility function, some combination of these factors, or something else entirely.

Research using structural modelling and *field* data has its own strengths and weaknesses. Field data reflect choices from real-world markets and hence have greater external validity than abstract and unfamiliar laboratory decisions. Research with field data can also take advantage of existing large datasets on household behavior. However, field data are difficult to interpret since the researcher does not know exactly what tradeoffs households actually face in the marketplace. Structural modelling helps to pin down some of these tradeoffs but such modelling relies on a large set of explicit and implicit assumptions.

Given all of these considerations, laboratory and field research complement each other. Hence it is surprising that efforts to estimate discount rates have primarily used laboratory evidence.[1] This imbalance is particularly true of the recent research on *generalized* time preferences (i.e., discount functions that are not restricted to the class of exponential functions). Hundreds of studies beginning with Chung and Herrnstein (1967) and reviewed in Ainslie (1992) and Frederick et al. (2002) have estimated generalized discount functions with laboratory evidence while only a handful have attempted to do this with field data.[2]

[1] Notable exceptions include Carroll and Samwick (1997), Gourinchas and Parker (2002), Hausman (1979), Lawrence (1991), Samwick (1998), Viscusi and Moore (1989), and Warner and Pleeter (2001).

[2] Prominent field studies include Attanasio and Weber (1995), Attanasio, Banks, Meghir and Weber (1999), Paserman (2004), Fang and Silverman (2004), and Shui and Ausubel (2004).

The current paper contributes to the literature estimating generalized discount functions using field data. We use lifecycle consumption data to estimate time preferences and to formally test the restriction to exponential discounting. Specifically, we use numerical methods to recursively solve and simulate a structural "buffer stock" model of lifecycle consumption and investment choices (Deaton 1991, Carroll 1992, Carroll 1997). Our model includes a rich array of financial instruments, constraints, demographic factors, and stochastic events – e.g., liquid and illiquid assets, revolving credit, liquidity constraints, household dependents, mortality, retirement, Social Security, and stochastic labor income – and thus controls for a number of relevant factors that affect intertemporal decisions.

We estimate the model's time preference parameters using a two-stage Method of Simulated Moments (MSM) procedure (McFadden 1989, Pakes and Pollard 1989, Duffie and Singleton 1993), which was first used to study lifecycle consumption behavior by Gourinchas and Parker (2002).[3] The MSM procedure extends the Generalized Method of Moments (GMM) to account for numerical simulation error. In the first stage of the MSM procedure we estimate inputs to the life-cycle model, including the parameters of the stochastic labor income process, interest rates, credit card borrowing limits, and parameters that describe variation in household size over the lifecycle. In the second stage of the MSM procedure we use the simulation model to estimate time preference parameters. These preference parameters are identified by empirical patterns of wealth accumulation, credit card borrowing, and consumption-income comovement. Uncertainty in estimates of the first stage parameters propagates to the standard errors for the time preference parameters estimated in the second stage. Formal incorporation of the first stage is critical since it raises our second-stage standard errors by nearly an order of magnitude.

Our analysis has three goals. First, this paper uses field data to estimate time preference parameters for households with a high school degree but not a college degree, comprising 59% of the US population. We study both the (restricted) exponential discount function and an (unrestricted) generalization that nests the exponential case. The unrestricted *quasi-hyperbolic* discount function

[3] Gourinchas and Parker (2002) and French (2005) use MSM to estimate different aspects of consumption models. Gourinchas and Parker identify the exponential discount function and the coefficient of relative risk aversion from lifecycle consumption profiles. French assesses how the opportunity to save and self-insure affects the impact of legislated Social Security and Medicare eligibility ages on the retirement decision. Most other applications of MSM have been in the industrial organization literature.

allows the discount rate to differ in the short-run and the long-run. Second, we formally test the restricted and unrestricted models, using both t-tests and overidentification tests. Finally, we ask whether these models accurately predict important empirical regularities in the lifecycle literature.

When we estimate a restricted (exponential) discount function, the MSM procedure estimates a single annual discount rate of 16.7%.[4] By contrast, when we estimate an unrestricted (quasi-hyperbolic) discount function, the MSM procedure estimates a short-run annualized discount rate of 39.5% and a long-run annualized discount rate of only 4.3%. All of these estimates are statistically significant at the 1% level. Our estimates also imply a formal rejection of the restricted case: i.e., we reject the hypothesis that the short-run discount rate is equal to the long-run discount rate.

Overidentification tests reinforce these conclusions. Only the exponential model is consistently rejected by overidentification tests. Intuitively, the exponential model cannot simultaneously explain high levels of credit card borrowing and high levels of retirement wealth accumulation. By contrast, the quasi-hyperbolic discount function implies that consumers will simultaneously act patiently and impatiently, because consumers have conflicting short-run and long-run discount rates.[5] In theory, low long-run discount rates explain why households accumulate substantial (illiquid) retirement wealth at real interest rates of about 5%, while high short-run discount rates imply that the *same* households borrow regularly on credit cards at real interest rates of 12%. By accumulating wealth in illiquid form, households commit themselves to act patiently in the future (i.e., not to spend down accumulated assets). However, when liquid assets and unused credit card balances are available, households spend when they can and therefore appear impatient.

We conclude the paper by reporting a wide range of robustness checks that reinforce some of our previous findings and identify the limits of our results. In 19 of 21 cases we find a significant gap between long-term and short-term discount rates, and in the other two cases we find a very large gap that is swamped by even larger standard errors. Quantitatively, our parameter estimates and standard errors are robust to many assumptions – e.g., about the persistence of the income process, the bequest motive, returns to scale within the household, and the credit card interest rate

[4]By contrast, most authors who calibrate exponential discount functions with lifecycle consumption and wealth data have adopted discount rates that are around 5% (Engen, Gale and Scholz 1994, Hubbard, Skinner and Zeldes 1994, Laibson, Repetto and Tobacman 1998, Engen, Gale and Uccello 1999). Our results differ because we ask our model to simultaneously fit wealth accumulation data *and* credit card borrowing data.

[5]See Angeletos, Laibson, Repetto, Tobacman and Weinberg (2001) and Laibson, Repetto and Tobacman (2003). These predictions rely in part on consumers' access to illiquid assets for long-run saving.

– but sensitive to assumptions about the return on illiquid assets and the coefficient of relative risk aversion.

We go further toward understanding the role of the coefficient of relative risk aversion by estimating it simultaneously with the discounting parameters. The MSM procedure results in estimates of the coefficient of relative risk aversion between 0.2 and 0.3 in both the exponential and quasi-hyperbolic cases. In the exponential case, the estimated discount rate falls, and the risk aversion coefficient is not well identified. In the quasi-hyperbolic case, the short-term and long-term discount rates also fall but continue to differ statistically, and all three parameters are estimated precisely.

This paper's findings are consistent with the results in other papers that have estimated quasi-hyperbolic time preference parameters with structural models and field data. Paserman (2004) obtains identification from heterogeneity in unemployment durations and reservation wages to find estimates of the short-run annualized discount rate that range from 11% to 91% and a long-run discount rate of only 0.1%. He rejects the exponential discounting null hypothesis for two of three subsamples. Fang and Silverman (2004) estimate models of both "sophisticated" and "naive" quasi-hyperbolic discounting.[6] Using data on welfare recipients, they find in the sophisticated case a short-run discount rate of 108% and a long-run discount rate of 13%, and they reject the null hypothesis of exponential discounting. Their results in the naive case are very similar. Finally, Shui and Ausubel (2004) use data from a direct mail credit card interest rate experiment to estimate the parameters of sophisticated and naive quasi-hyperbolic models. They obtain short-run discount rates of 24% in the sophisticated case and 20% in the naive case. In both cases they find a long-term discount rate of 0.01% and they reject the exponential discounting null hypothesis.

The empirical data we use to estimate our model are presented in Section 2. Section 3 summarizes the structural model that we use. We explain the MSM procedure in Section 4. Section 5 presents our results. Section 6 discusses extensions and Section 7 concludes.

[6] Naive hyperbolic decision-makers incorrectly believe that they will have exponential discount functions in the future, while hyperbolics who are aware that they will be hyperbolic in the future are called "sophisticates." See Akerlof (1991) and O'Donoghue and Rabin (1999a, 1999b) for analysis of naive hyperbolic discounters. In most consumption models sophisticates and naifs behave similarly (Angeletos et al. 2001). We focus on the sophisticated case. See Section 6 for more discussion.

2 Wealth Accumulation, Credit Card Borrowing, and Consumption-Income Comovement Data

We estimate exponential and quasi-hyperbolic discount functions by matching moments that characterize wealth accumulation, credit card borrowing, and excess sensitivity of consumption to predictable movements in income. We summarize these statistics in this section. Table 1 reports these moments, and Appendix A contains a detailed description of the data sources and estimation procedures. All of the analysis that we conduct applies only to U.S. households whose head has a high school degree but not a college degree. These households constitute 59% of the population (U.S. Census Bureau 1995).[7]

The first statistic, % $Visa$, is the fraction of households that borrow on credit cards.[8] Our analysis finds that 67.8% of households *pay interest* on credit card debt in any given month. Specifically, % $Visa$ represents the fraction of households that self-report that they did not pay their bill in full at the end of the last month (SCF 1995 and 1998). Though there is considerable heterogeneity among households, credit card borrowing is ubiquitous across the entire distribution of wealth. Table 2 of Laibson et al. (2003) reports the fraction of households borrowing on credit cards by age and by wealth quartile.[9] Among households with a head between ages 20-29 that are in the top wealth quartile for their age group, three-fourths did not repay their credit card bills in full the last time they paid their bills. For households with a head in his or her thirties, over 80% of median wealth-holders had credit card debt. Even among the households with a head between ages 50-59 that are between the 50th and 75th wealth percentiles, 56% borrowed and paid interest on credit card debt in the past month. The typical American household accumulates wealth in the years leading up to retirement and simultaneously borrows on their credit cards.

We construct the second statistic, *mean* $Visa$, by dividing credit card borrowing by mean age-specific income.[10] We then average this fraction over the lifecycle. The average household has

[7] Laibson et al. (1998, 2003) examine households in *all* education categories using a calibration framework instead of an estimation framework. They find qualitatively similar results across education categories.

[8] This is the fraction that borrows on any type of card, not just Visa cards.

[9] Note however that the lifecycle patterns in this table are difficult to interpret because they reflect both cohort effects and age effects.

[10] Throughout the paper we exclude households reporting less than $1000 in annual income. Since our income measure includes government and inter-household transfers, we view income less than $1000 as having a high chance of reflecting mismeasurement.

outstanding credit card debt equal to 11.7% of the mean income of its age cohort (SCF and Federal Reserve Board, 1995 and 1998). This statistic is effectively a ratio of means. We use the ratio of means instead of the mean ratio, since household level income can take small values.

The third statistic, CY, represents the excess sensitivity of consumption in response to predictable income changes.[11] We estimate that the marginal propensity to consume is 23% of the expected income change (PSID 1978-1992). This figure is consistent with other analyses in the literature.[12]

The final statistic, *wealth*, is a weighted average of wealth-to-income ratios across households with heads aged 50-59, *excluding* 'involuntary' wealth like Social Security and other defined benefit pensions. We restrict attention to households aged 50-59 so that we are primarily measuring retirement savings. We also weight the wealth-to-income ratios with a scaled arctan function to downweight (positive and negative) outliers in wealth and to simultaneously prevent low values in income from causing the statistic to blow up.[13] The resulting *wealth* measure equals 2.60 (SCF 1983-1998). For comparison, the median wealth-to-income ratio for the same sub-sample is 2.16.[14]

3 Consumption-Savings Model

Our work extends the numerical simulation literature pioneered by Carroll (1992, 1997), Deaton (1991), Zeldes (1989), Gourinchas and Parker (2002), and Hubbard, Skinner and Zeldes (1994, 1995). Our specific analysis is most closely related to Gourinchas and Parker (2002) — we adopt their MSM procedure — and Laibson, Repetto and Tobacman (2003) — we adopt their calibrated structural model. In the next section we review the Gourinchas and Parker MSM procedure. In the current section we review the Laibson, Repetto, and Tobacman model.

The parameters for this structural model can be found in Table 2. In the model, economic de-

[11] See for example Carroll and Summers (1991), Shea (1995), and Parker (1999).

[12] Most previous work on excess sensitivity has found coefficients between 0 and 0.5. See Deaton (1992) and Browning and Lusardi (1996) for reviews.

[13] The smoothness of arctan (in contrast to, say, the median) implies requisite differentiability of the theoretical moment conditions. Our results are robust to different choices of this transformation. See Appendix A for details on the scaling.

[14] We could not use an easy-to-interpret ratio of means to construct the *wealth* moment (as we did for $mean\ Visa$) because wealth is so strongly skewed to the right. For instance, the ratio of mean to median is six times greater for wealth than for income (SCF, 1998): (mean income)/(median income) = 1.9; (mean wealth)/(median wealth) = 11.8.

cision making begins at age 20. Households have an age-dependent survival hazard of s_t calibrated with data from the U.S. National Center for Health Statistics (1994). Household composition varies deterministically with age as children and adult dependents enter and leave the household.[15] Effective household size n_t equals the number of spouses – which we assume to be two in our benchmark model – plus the number of dependent adults, plus 0.4 times the number of children under 18. In our benchmark analysis we assume that the spouses die simultaneously, but we relax this assumption in Section 5.3.3 of our robustness analysis, where we also analyze different assumptions for the returns to scale in household consumption.

Let Y_t represent period t after-tax income from transfers and wages, including labor income, inheritances, private defined benefit pensions, and government transfers including Social Security. During working life $y_t = \ln(Y_t)$ is modelled as the sum of a cubic polynomial in age, an AR(1), and an iid shock. We approximate the AR(1) with a Markov process, and denote the Markov state ζ. During retirement, y_t is the sum of a linear polynomial in age and an iid shock. Retirement occurs exogenously at age T. The income process and the retirement age are calibrated from the PSID.[16]

Let X_t represent liquid asset holdings at the beginning of period t before receipt of Y_t. If $X_t < 0$ then uncollateralized debt — i.e. credit card debt — was held between $t-1$ and t. Households face a credit limit at age t of λ times average income at age t, i.e., $X_t \geq -\lambda \overline{Y}_t$.[17] The model precludes consumers from simultaneously holding liquid assets and credit card debt, though such potentially suboptimal behavior has been documented among a subpopulation of consumers by Gross and Souleles (2002a) and Bertaut and Haliassos (2001).

Positive liquid asset holdings earn a risk-free real after-tax gross interest rate of R, the average of Moody's AAA municipal bond yields from 1980-2000 (Gourinchas and Parker 2002). Households pay a gross "effective" real interest rate on credit-card borrowing of R^{CC}. We refer to this simply as the credit card interest rate, but our estimate of R^{CC} captures the impact of bankruptcy and

[15] The demographic profiles are estimated parametrically using the PSID.

[16] Some authors, such as Hubbard et al. (1994), find very similar parameter values. Other authors estimate more persistent income shocks than we find. Finally, some papers estimate non-stationary processes. Overidentification tests fail to reject our results, and the discounting estimates we report below are robust to increasing the persistence of the income process. See Appendix Table 1.

[17] The limit is estimated from the SCF. This is a crude representation of the income-based credit limits that are common in the revolving credit market. Assets are not an important determinant of credit card borrowing limits because large asset classes like retirement accounts (and in some states home equity) can not be seized after a credit card default.

inflation, which lower consumers' effective interest payments.[18]

Let Z_t represent (net) illiquid asset holdings at the beginning of period t, with $Z_t \geq 0, \forall t$. Illiquid assets include durables, which generate two types of returns: capital gains and consumption flows. For computational tractability, capital gains equal zero (i.e., $R^Z = 1$) and the annual consumption flow is $\gamma Z_t = 0.05 \cdot Z_t$. Hence, the return from holding the illiquid asset is a 5% annual flow of consumption. We also adopt the assumption that Z can only rise during the owner's lifetime; i.e., transaction costs are large enough that the Z asset is never sold until wealth is bequeathed to the next generation. These choices about Z do not match the properties of a particular illiquid asset though Z has some of the features of home equity.[19]

Four observations motivate these assumptions about Z. First, despite increasing financial sophistication many household assets continue to be partially illiquid and were certainly illiquid during our sample period (1978 to 1998). Accessing equity in homes, cars and retirement plans like 401(k)s entails at least small transactions costs and delays. Second, theory (Laibson 1997) and simulations (Laibson et al. 2003) have shown that for the quasi-hyperbolic case, small transactions costs have the same impact on consumers as complete illiquidity. Third, illiquidity of Z mimics the optimal savings mechanisms that have recently been derived for quasi-hyperbolic consumers (Amador, Werning and Angeletos 2004). Finally, assuming that Z is illiquid increases computational tractability by limiting the choice set of the consumer. In Subsection 5.3 we evaluate the robustness of our set-up by making Z more attractive, and we highlight the potential drawbacks of assuming that Z is illiquid when we discuss extensions of the model in Subsection 6.4.

---

[18] Specifically, from the quarterly interest rates reported in the Federal Reserve Board's G-19 historical series, we subtract the CPI-U and the bankruptcy rate. We calculate the latter by dividing the number of bankruptcies (American Bankruptcy Institute) by the number of US households that have credit cards (SCF). This attributes all US bankruptcies to households holding credit cards.

These calculations omit two considerations that would further lower the effective credit card rate. First, without declaring bankruptcy, households might be able to default on their credit card debt. Second, consumption may be unusually low in the bankruptcy/default state, which reduces the cost of borrowing since repayment only occurs in "good states." On the other hand, we also omitted two bankruptcy-related considerations that raise the effective credit card rate. The model does not account for the stigma associated with bankruptcy (Gross and Souleles 2002b) or for the cost of future exclusion from credit markets. Robustness checks — using a wide range of credit card interest rates — are provided in section 5.3.

[19] Consider a consumer who owns a house of fixed real value H and derives annual consumption flows from the house of γH. Suppose the consumer has a mortgage of size M and home equity of $H - M$. The real cost of the mortgage is ηM, where $\eta = i \cdot (1 - \tau) - \pi$ is the nominal mortgage interest rate adjusted for inflation and the tax deductibility of interest payments. If we assume $\eta \approx \gamma$, the net benefit to the homeowner is $\gamma H - \eta M \approx \gamma(H - M) = \gamma Z$. Section 6 discusses enriching the modelling of Z.

Let I_t^X represent net investment into the liquid asset X during period t, and let I_t^Z represent net investment into the illiquid asset Z during period t. Then the dynamic budget constraints are given by,

$$X_{t+1} = R^X \left(X_t + I_t^X\right) \quad (1)$$

$$Z_{t+1} = R^Z \left(Z_t + I_t^Z\right). \quad (2)$$

Since the interest rate on liquid wealth R^X depends on whether the consumer is borrowing or saving in her liquid accounts,

$$R^X = \begin{cases} R^{CC} & \text{if} \quad X_t + I_t^X < 0 \\ R & \text{if} \quad X_t + I_t^X \geq 0 \end{cases}$$

Denote $r^{CC} = R^{CC} - 1$. The static budget constraint is:

$$C_t = Y_t - I_t^X - I_t^Z$$

The state variables Λ_t at the beginning of period are age (t), liquid wealth ($X_t + Y_t$), illiquid wealth (Z_t), and the value of the Markov process (ζ_t). The non-redundant choice variables are I_t^X and I_t^Z. Consumption is calculated as a residual.

The consumer has constant relative risk aversion and a quasi-hyperbolic discount function. For $t \in \{20, 21, ..., 90\}$, self t has instantaneous payoff function

$$u(C_t, Z_t, n_t) = n_t \cdot \frac{\left(\frac{C_t + \gamma Z_t}{n_t}\right)^{1-\rho} - 1}{1 - \rho}$$

and continuation payoffs given by

$$\beta \sum_{i=1}^{90-t} \delta^i \left(\Pi_{j=1}^{i-1} s_{t+j}\right) \left[s_{t+i} \cdot u(C_{t+i}, Z_{t+i}, n_{t+i}) + (1 - s_{t+i}) \cdot B(X_{t+i}, Z_{t+i})\right].$$

Here ρ is the coefficient of relative risk aversion, and $B(\cdot)$ represents the payoff in the death state, which incorporates a bequest motive.[20] The first expression in the bracketed term is the utility

[20] Liquidated bequeathed wealth is consumed by heirs as an annuity (Laibson et al. 2003). Specifically, if $\bar{n}$

flow that arises in period $t+i$ if the household survives to age $t+i$. The second expression is the termination payoff in period $t+i$ which arises if the household dies between period $t+i-1$ and $t+i$. The quasi-hyperbolic discount function $\{1, \beta\delta, \beta\delta^2, \beta\delta^3, ...\}$ corresponds to a short-run discount rate of $-\ln(\beta\delta)$ and a long-run discount rate of $-\ln(\delta)$.[21] When $\beta = 1$ the consumer has exponential discounting, which implies that the agent is dynamically consistent.

Following Strotz (1955) we model behavior as an intra-personal game among selves $\{20, 21, ..., 90\}$. Taking the strategies of other selves as given, self t picks a strategy at time t. This strategy is a map from the Markov state $\Lambda_t = \{t, X+Y, Z, \zeta\}$ to the choice variables $\{I^X, I^Z\}$. An equilibrium is a fixed point in the strategy space, such that all strategies are optimal given the strategies of other players. We solve for equilibrium strategies using numerical backwards induction.

Let $V_{t,t+1}(\Lambda_{t+1})$ represent the time $t+1$ continuation payoff function of self t. Then self t's objective function is

$$u(C_t, Z_t, n_t) + \beta\delta E_t V_{t,t+1}(\Lambda_{t+1}). \tag{3}$$

Self t chooses $\{I^X, I^Z\}$ in state Λ_t to maximize this expression. The sequence of continuation payoff functions is defined recursively

$$V_{t-1,t}(\Lambda_t) = s_t[u(C_t, Z_t, n_t) + \delta E_t V_{t,t+1}(\Lambda_{t+1})] + (1-s_t)E_t B(\Lambda_t). \tag{4}$$

The induction continues in this way.

We generate $J_s = 5000$ independent streams of income realizations for J_s households. Then we simulate lifecycle choices for these households, assuming they make equilibrium decisions conditional on their state variables. From the simulated profiles of C, X, Z, and Y, we calculate the moments used in the MSM estimation procedure. Note that the simulated profiles, and hence the summary

is average effective household size over the life-cycle, $\bar{y}$ is average labor income over the life-cycle, and $u_1(\bar{y}, 0, \bar{n})$ is the partial derivative of instantaneous utility $u(C, Z, n)$ with respect to consumption, we assume, $B(X, Z) = (R-1) \cdot \max\left\{0, X + \frac{2}{3}Z\right\} \cdot \frac{u_1(\bar{y}, 0, \bar{n})}{1-\delta}$. Often liquiditating bequeathed wealth entails large transactions costs, so we multiply bequeathed illiquid wealth by two-thirds. The rest of the expression follows because the total consumption of the bequest recipient approximately equals $\bar{y}$, and on average the effective household size of the bequest recipient equals $\bar{n}$.

This formulation is consistent with both common assumptions in the literature on bequests: it is a parametrization of an altruistic bequest motive which also assumes a "warm-glow"-style payoff linear in the size of the bequest. Adjusting the bequest payoff by 25% has a trivial impact on our parameter estimates. See Appendix Table 1.

[21] See Phelps and Pollak (1968) and Laibson (1997).

moments, depend on the parameters of the model. Since the model cannot be solved analytically, its quantitative predictions are derived from the simulated lifecycle profiles. Variability arising from (finite sample) simulation error is addressed in the estimation procedure.

4 Method of Simulated Moments Procedure

We estimate the parameters of the model's discount function in the second stage of a Method of Simulated Moments procedure, closely following the methodology of Gourinchas and Parker (2002). MSM allows us to evaluate the predictions of our model, to formally test the nested null hypothesis of exponential discounting, $\beta = 1$, and to perform specification tests. We use MSM rather than GMM because the model cannot be solved analytically.[22] The current section describes our procedure. Appendix B presents derivations and some technical details.

Our MSM procedure has two stages. In the first stage, nuisance parameters, χ, are estimated using standard GMM techniques (see Table 2). We take these $N_\chi = 28$ estimated parameters and their associated variances, Ω_χ, from Laibson et al. (2003).[23] Some authors describe this first-stage as the "calibration" stage. These first-stage estimates match those of numerous other researchers.[24]

Given $\hat{\chi}$ and Ω_χ, the second stage uses additional data and more of the model's structure to estimate N_θ additional parameters θ.[25] The second stage, taking the first stage parameters fixed at $\hat{\chi}$, chooses θ to minimize the distance between the empirical and the simulated moments.

[22] See McFadden (1989), Pakes and Pollard (1989), and Duffie and Singleton (1993) for the first formulations of MSM, and Stern (1997) for a review of simulation-based estimation techniques.

[23] Included in χ are seven pre-retirement income level coefficients, three pre-retirement income variability coefficients, the retirement age, five post-retirement income coefficients, one post-retirement income variability coefficient, six effective household size coefficients, the credit limit, the coefficient of relative risk aversion, and three interest rates.

[24] Hubbard et al. (1994) report an almost identical process for after-tax non-asset income. Attanasio and Weber (1995) report similar family size profiles over the life cycle. Bernheim, Skinner and Weinberg (2001) provide similar estimates of the retirement age and income replacement rates. Ausubel (1991) reports credit card interest rates that are similar to our estimates before we correct for personal bankruptcy.

[25] In principle θ and χ could be estimated simultaneously. In practice, computational considerations led us to adopt the two stage estimator, which has already been successfully employed in the household finance literature (Gourinchas and Parker 2002). The consistency of our (second stage) θ estimator depends on strong exogeneity assumptions. For example, if labor supply were instead endogenously chosen, then households could insure idiosyncratic shocks by working more hours or postponing retirement. Introducing such features into the model would make credit card borrowing more puzzling, since endogenous labor supply could be used instead of credit card borrowing to smooth consumption during transitory periods of low wages. Hence, we conjecture that our assumption of exogenous labor supply is biasing up our estimates of β. However, our exogeneity assumptions may induce other offsetting biases, and their overall effect on the estimates of β is unclear.

Specifically, we use the data from Section 2 on wealth accumulation, credit card borrowing, and excess sensitivity to estimate $\theta = (\beta, \delta)$ in the second stage. MSM differs from a calibration exercise followed by a one-stage estimation in that it propagates uncertainty in the first stage parameters into the standard errors of the second stage parameter estimates. In other words Ω_θ, the variance matrix of $\hat{\theta}$, depends on Ω_χ.[26] For three of the model's parameters that are not pinned down precisely by available data, r^{CC}, γ, and ρ, we perform additional robustness checks in Subsection 5.3.

Denote the empirical vector of N_m second stage aggregate moments by $\bar{m}_{J_m}$. Let J_m be the numbers of empirical observations used to calculate the elements of $\bar{m}_{J_m}$.[27] Denote the theoretical population analogue to $\bar{m}_{J_m}$ by $m(\theta, \chi)$ and let $m_{J_s}(\theta, \chi)$ be the simulation approximation to $m(\theta, \chi)$. Let $g(\theta, \chi) \equiv [m(\theta, \chi) - \bar{m}_{J_m}]$ and $g_{J_s}(\theta, \chi) \equiv [m_{J_s}(\theta, \chi) - \bar{m}_{J_m}]$. The moment conditions imply that in expectation

$$Eg(\theta_0, \chi_0) = E[m(\theta_0, \chi_0) - \bar{m}_{J_m}] = 0,$$

where (θ_0, χ_0) is the true parameter vector. Define derivatives of the moment functions with respect to the parameters by $G_\theta \equiv \frac{\partial g(\theta_0, \chi_0)}{\partial \theta}$ and $G_\chi \equiv \frac{\partial g(\theta_0, \chi_0)}{\partial \chi}$. Let Σ_g be the variance-covariance matrix of the second stage moments in the population. Let $\Omega_g \equiv E\left[g(\theta_0, \chi_0) g(\theta_0, \chi_0)'\right]$ be the variance of the second stage moment estimates $\bar{m}_{J_m}$, which is estimated directly and consistently from sample data.[28]

Let W be a positive definite N_mxN_m weighting matrix. Define

$$q(\theta, \chi) \equiv g_{J_s}(\theta, \chi) \cdot W^{-1} \cdot g_{J_s}(\theta, \chi)' \tag{5}$$

[26] Our derivation of Ω_θ assumes that the first-stage moments and the second-stage moments have uncorrelated measurement error. We made this simplifying assumption because most of the data we use to identify θ and χ come from separate datasets. Exceptions are the credit limit and CY. However, even household level covariances between the second stage moments and the credit limit are approximately zero. Moreover, CY's large standard error means that it is has little weight in our second-stage estimation anyway – our results do not qualitatively change when we omit CY from the list of moments.

[27] Though the main text does not discuss it, the procedure accounts for the fact that J_m differs for different moments. Appendix B contains details.

[28] If the same number of empirical observations $\bar{J}_m$ were available to calculate all of the second stage moments, then we would have $\Omega_g = \Sigma_g / \bar{J}_m$.

as a scalar-valued loss function, equal to the weighted sum of squared deviations of simulated moments from their corresponding empirical values. Then our procedure is to fix χ at the value of its consistent first-stage estimator, minimize the loss function $q(\theta, \hat{\chi})$ with respect to θ, and define the estimator as[29]

$$\hat{\theta} = \arg \min_{\theta} \quad q(\theta, \hat{\chi}). \tag{6}$$

Pakes and Pollard (1989) demonstrate that under regularity conditions satisfied here $\hat{\theta}$ is a consistent estimator of θ_0, and $\hat{\theta}$ is asymptotically normally distributed. As shown in Appendix B,

$$\Omega_\theta = Var\left(\hat{\theta}\right) = \left(G'_\theta W G_\theta\right)^{-1} G'_\theta W \left[\Omega_g + \Omega_g^s + G_\chi \Omega_\chi G'_\chi\right] W G_\theta \left(G'_\theta W G_\theta\right)^{-1}, \tag{7}$$

where $\Omega_g^s = \frac{J_m}{J_s}\Omega_g$ is the simulation correction.

This equation is used to calculate standard errors for our estimates of θ. All derivatives are replaced with consistent numerical analogues, which we calculate using the model and simulation procedure.[30] We estimate Ω_g and Ω_χ consistently from sample data. After obtaining estimates using the weighting matrix $W = \Omega_g^{-1}$, we can construct the optimal weighting matrix $W_{opt} = \left[\Omega_g + \Omega_g^s + G_\chi \Omega_\chi G'_\chi\right]^{-1}$. Many authors (Altonji and Segal 1996, West, Wong and Anatolyev 2004, for example) have found optimally-weighted GMM procedures lead to biased estimates in small samples, so our baseline estimates use the simple weighting matrix $W = \Omega_g^{-1}$. In robustness checks we find that our qualitative conclusions are not affected by adoption of either of these weighting matrices.

To interpret the expression for Ω_θ, first consider the simulation correction Ω_g^s. As the size of the simulated population J_s relative to the size of the sample J_m goes to infinity, the simulation correction approaches zero. Intuitively, as the simulation becomes an ideal approximation for the true population, the simulation correction disappears. Next examine the first stage correction $G_\chi \Omega_\chi G'_\chi$. This correction increases with the uncertainty Ω_χ in our estimates of the first-stage parameters; note that Ω_χ itself is increasing in the underlying population variance of χ and decreasing

[29] We perform this minimization with Matlab's Nelder-Mead simplex algorithm. This algorithm is slower but more robust than derivative-based methods, and here it is preferred because of the nonconvexities in quasi-hyperbolic policy functions.

[30] We take numerical derivatives on both sides of the optimum and accept the derivative that has the most conservative implications for Ω_θ.

in the number of observations we use to estimate χ. The first-stage correction also increases with the sensitivity of the second-stage moments to changes in the first-stage parameters.

When neither the simulation correction nor the first stage correction matter, we obtain,

$$\Omega_\theta = \left(G'_\theta W G_\theta\right)^{-1} G'_\theta W \Omega_g W G_\theta \left(G'_\theta W G_\theta\right)^{-1}.$$

In the benchmark case where we assume $W = \Omega_g^{-1}$, this becomes the standard GMM variance formula: $\Omega_\theta = \left(G'_\theta \Omega_g^{-1} G_\theta\right)^{-1}$.

MSM also allows us to perform specification tests. If the model is correct,

$$\begin{aligned} \xi\left(\hat{\theta}, \hat{\chi}\right) &\equiv g_{J_s}\left(\hat{\theta}, \hat{\chi}\right) \cdot W_{opt} \cdot g'_{J_s}\left(\hat{\theta}, \hat{\chi}\right) \\ &= g_{J_s}\left(\hat{\theta}, \hat{\chi}\right) \cdot \left[\Omega_g + \Omega_g^s + G_\chi \Omega_\chi G'_\chi\right]^{-1} \cdot g'_{J_s}\left(\hat{\theta}, \hat{\chi}\right) \end{aligned}$$

will have a chi-squared distribution with $N_m - N_\theta$ degrees of freedom. This test statistic equals $q\left(\hat{\theta}, \hat{\chi}\right)$ in the optimal-weighting case.

5 Results

In this section we discuss the paper's three sets of findings. We report estimates for the discount factors β and δ, including the special case in which we impose $\beta = 1$ (leaving δ as the only free parameter). Second, we evaluate the statistical fit of the estimated models, using both t-tests and overidentification tests. Finally, we ask whether these models accurately predict key empirical regularities in the lifecycle literature.

The coefficient of relative risk aversion ρ, the return on illiquid assets γ, and the credit card interest rate r^{CC} affect the quantitative results and these parameters are difficult to pin down empirically. As benchmarks we adopt the assumptions $\gamma = 5\%$, $r^{CC} = 11.52\%$, and $\rho = 2$, but we also examine the robustness of our findings to changes in these parameters in subsection 5.3, including the simultaneous estimation of β, δ and ρ. We also discuss the sensitivity of our results to alternative assumptions on the returns to scale in household consumption and on household mortality.[31] Unless otherwise specified we study estimates based on the robust weighting matrix $W = \Omega_g^{-1}$, but we include some representative results using the efficient weighting matrix.

5.1 Identification

Identification of β and δ depends on the way the simulated moments $m_{J_s}(\theta, \hat{\chi})$ vary as functions of β and δ. In our case the pre-retirement weighted wealth-to-income ratio, *wealth*, increases in both β and δ. All of the other moments decrease in β and δ. Hence, β and δ are substitutes. Nevertheless, the model is identified since β and δ are not *perfect* substitutes.

Identification of β and δ can be visualized in Figure 1, which plots $q(\theta, \hat{\chi})$. Recall that q is a weighted sum of squared deviations of the simulated moments from their empirical analogs. Smaller values of q reflect a closer fit between the simulated model and the data. In Figure 1, q resembles an upward-opening paraboloid. Intermediate values of β and δ — neither zero nor one — minimize the distance between the simulated moments and the empirical moments.

Figure 1 exhibits an extended valley in the plot of q, traversing from high δ and low β to low δ and high β. The orientation of this valley implies that β and δ are partial substitutes; when δ is high, low values of β best match the empirical facts, and vice-versa. If the valley had a flat bottom, β and δ would be perfect substitutes and the model would not be identified.

The lowest point in the valley is $(\beta, \delta) = (0.703,\ 0.958)$, and the paraboloid rises steeply as β rises. As β approaches 1 the model does very poorly; near $\beta = 1$, matching credit card data requires a low value of δ (e.g., $\delta \approx 0.85$), but with δ this low, wealth accumulation vanishes. Figure 2 displays a higher-resolution plot, which highlights the fact that the model cannot match the data when β is close to 1.

The figures reflect the intuition that *low* long-term discount rates are necessary to match observed levels of retirement wealth. A household will only accumulate illiquid wealth that has a return of about 5% if the household's long-term discount rate is not much greater than 5%. If the long-term discount rate is pinned down in this way, it is necessary to have a β value below one to match the data on credit card borrowing.

[31] Appendix Table 1 reports the results of adjusting the bequest payoff and of increasing the persistence of income shocks.

5.2 Benchmark Estimates

We report our benchmark estimates in Table 3. In the unconstrained case (Column 1) the MSM procedure yields an estimate of $\hat{\beta} = 0.703$, with a standard error (s.e.(i) in the Table) of 0.109. For this specification $\hat{\beta}$ lies significantly below 1; the t-stat for the $\beta = 1$ hypothesis test is $t = \frac{1-0.703}{0.109} = 2.72$. The MSM procedure yields an estimate of $\hat{\delta} = 0.958$, with a standard error of 0.007. The estimated values of β and δ imply a short-run discount rate of $-\ln(0.703 * 0.958) = 39.5\%$ and a long-run discount rate of $-\ln(0.958) = 4.3\%$.

At the estimated parameter values, the quasi-hyperbolic model generates the moment predictions reported in Column 1 of the lower panel of Table 3. We can compare these simulated moments with the sample moments $\bar{m}_{J_m}$, which are reproduced in Column 5. Qualitatively, the model predicts both active borrowing on credit cards and accumulation of midlife wealth. Quantitatively, the model predicts a fraction borrowing three standard errors from the sample value, a level of borrowing that differs by five standard errors, and a consumption-income comovement coefficient and measure of wealth accumulation that are both off by about one standard error. However, such comparisons do not account for uncertainty in the first-stage parameters of the model. Once this uncertainty is propagated through the estimation procedure — which is what MSM does — the model is consistent with the data. The (inverse) goodness-of-fit measure $\xi\left(\hat{\theta},\, \hat{\chi}\right) = 3.01$ compares favorably to the 5% critical value of 5.99 for a chi-squared distribution with two degrees of freedom. For the benchmark case, we fail to reject the overidentification test.

We also estimate δ alone, imposing the restriction $\beta = 1$. This exponential discounting case yields the results in Column 2. We find $\hat{\delta} = 0.846$ – implying a discount rate of 16.7% – and a standard error of 0.025. At these point estimates, the empirical facts about credit card borrowing and excess sensitivity are matched well. However, with such a high discount rate the model cannot account for observed wealth data. Instead, it predicts $wealth = -0.05$; *wealth* loses in the tug of war between fitting *wealth*, which requires a low discount rate, and fitting the credit card variables $\%\ Visa$ and $mean\ Visa$, which requires a high discount rate.[32] The best fit avail-

[32] The empirical value of *wealth* is 2.6, twenty standard errors from its simulated value of $wealth\left(\hat{\theta}, \hat{\chi}\right) = -0.05$. However, matching the empirical value of *wealth* would require $\%\ Visa$ to approach 0, forty standard errors from its empirical value. If the returns to illiquid wealth (i.e., γ) were high enough, an exponential model could more successfully match the facts simultaneously. The results from Case B in Subsection 5.3.4 provide suggestive evidence.

able under an exponential model predicts that typical households have negative total assets in their peak accumulation years. Goodness of fit naturally suffers: $\xi(\beta = 1,\ \delta = 0.846,\ \hat{\chi}) = 217 \gg 3.01 = \xi(\beta = 0.703,\ \delta = 0.958,\ \hat{\chi})$. With the exponential restriction we estimate only one parameter, but $\xi(\beta = 1,\ \delta = 0.846,\ \hat{\chi})$ compares unfavorably to even the 1% critical value of 11.34 for a chi-squared distribution with three degrees of freedom. Recall that above we compared $\xi(\beta = 0.703,\ \delta = 0.958,\ \hat{\chi})$ to a chi-squared distribution with only two degrees of freedom. This difference accounts for the degree of improvement in goodness-of-fit possible merely by adding a free parameter. The p-value represents the smallest significance level at which the benchmark model would be rejected, so the overidentification test rejects the exponential model at the 1% level. The p-value for the unconstrained model exceeds the p-value for the constrained model by many orders of magnitude.

The standard errors reported as "s.e.(i)" in Table 3 incorporate corrections for the first stage estimation and for the simulation error. For comparison, we also report standard errors without these corrections: s.e.(ii) only includes the first stage correction, s.e.(iii) only includes the simulation correction, and s.e.(iv) includes neither.

Comparing s.e. (i) and (ii) reveals that if the simulation were infinitely large, so that the simulation exactly captured the properties of the theoretical population, the standard error on β would barely change – falling from 0.1093 to 0.1090. However, the standard errors are *dramatically* affected by the first stage correction. Comparing s.e. (i) and (iii), if the first stage parameters were known with certainty the standard error on β would shrink by a factor of six – falling from 0.1093 to 0.0170.

Using the optimal weighting matrix largely preserves the pattern of the benchmark results. Our optimal-weights findings are reported in Columns 3 and 4. The quasi-hyperbolic results with the optimal weighting matrix are similar to those with $W = \Omega_g^{-1}$. The estimated $\hat{\beta}$ and $\hat{\delta}$ are slightly higher, the standard error on $\hat{\beta}$ is lower, and the standard error on $\hat{\delta}$ is higher. In the exponential case, $\hat{\delta}$ is found to be substantially larger than in the benchmark; now $\hat{\delta}$ is selected by the estimation procedure to match the data on wealth at the expense of matching the facts on borrowing and excess sensitivity.

Uncertainty in all of the first stage parameters except γ, r^{CC}, and ρ has been incorporated into the standard errors reported above.[33] However, γ, r^{CC}, and ρ are difficult to pin down empirically. In the next subsection we explore the robustness of our findings to changes in those parameters and other components of the model.

5.3 Robustness

5.3.1 Interest Rate Perturbations

We perturb the return on the illiquid asset (γ) and the credit card interest rate (r^{CC}) one-by-one from their benchmark values (i.e., $\gamma = 5\%$ and $r^{CC} = 11.52\%$). We report the resulting estimates of β and δ in Columns 2-5 of Table 4A.

In Column 1 we reproduce the benchmark results as a reference. In Column 2 we set $\gamma = 3.38\%$, corresponding to the average tax- and inflation-adjusted mortgage interest rate from 1980-2000, as calculated from Freddie Mac's historical series of nominal mortgage interest rates and the CPI-U, assuming a marginal tax rate of 25%. Intuitively, this choice for γ reflects the interest savings resulting from paying off a dollar of mortgage debt. We interpret 3.38% as being at the low end of a range of possible assumptions about returns to the illiquid asset Z. In Column 2 we estimate a lower $\hat{\beta}$ and a higher $\hat{\delta}$ than in the benchmark case. Intuitively, when the return on the illiquid asset is relatively low, consumers will only invest in the illiquid asset if they have a low long-run discount rate. Lowering the long-run discount rate makes the consumer more patient generally, so the MSM procedure generates a greater short-run discount rate (i.e., a lower value of β) to offset this effect and thereby maintain high rates of credit card borrowing and consumption-income comovement. As a net result of these changes, the over-identifying restrictions are now rejected.

In Column 3, we consider the case $\gamma = 6.59\%$. Flavin and Yamashita (2002) calculate this as the *average* real after-tax return to housing, including capital gains, use-value, maintenance costs, and taxes.[34] We view this figure as falling toward the upper end of a range of possible returns to the illiquid asset in our model. If our model distinguished between average and marginal returns we would want to use the marginal return not the average return on housing.[35] Moreover, even the

[33] Our measure of r^{CC} is constructed from aggregate data, so its true variability is underestimated in the first stage.

[34] Flavin and Yamashita (2002) assume that use-value is equal to the real return on financial assets plus housing depreciation plus property taxes.

[35] Because of transaction costs, at the margin it may be optimal to pay off the mortgage rather than buying a larger house. The marginal return to paying off a mortgage is 3.38% as we argue above.

average return to housing may be lower than 6.59% because of transaction costs of buying/selling real estate.

Using $\gamma = 6.59\%$ we obtain a higher estimate of $\hat{\beta}$ and a lower estimate for $\hat{\delta}$ than in the benchmark case. As γ approaches the credit card interest rate, the model can more easily accommodate simultaneous illiquid wealth accumulation and credit card borrowing. Despite the increased estimate for $\hat{\beta}$, the $\beta = 1$ hypothesis is still rejected at the 99% confidence level. We also report a borderline rejection of the over-identifying restrictions for this case.

The large effects resulting from changes in γ contrast with the small effects we now report arising from changes in the credit card real interest rate, r^{CC}. In Column 4 we assume $r^{CC} = 10\%$ and find that $\hat{\beta}$ rises and $\hat{\delta}$ falls relative to the benchmark case. Column 5 shows similar effects in the opposite direction for $r^{CC} = 13\%$. We introduce these perturbations for two reasons. First, formal incorporation of uncertainty in r^{CC} through the first stage correction only accounts for variation in population average interest rates. Additional tests in Columns 4 and 5 could capture individual-level variation. In addition, these changes correspond to different perspectives on how bankruptcy matters for the cost of credit card borrowing. Our benchmark value for r^{CC} equals the debt-weighted average interest rate from the Fed, minus inflation, minus the personal bankruptcy rate. This ignores the fact (i) that the marginal utility of consumption may be especially high in the bankruptcy state, and that households may default without declaring bankruptcy, implying that our correction is too small, and (ii) that bankruptcy carries stigma (Gross and Souleles 2002b), implying that our correction is too large. We favor the middle specification as our benchmark, and note that changes in r^{CC} of about 150 basis points reported in Columns 4 and 5 have little quantitative effect on the time preference estimates. However, these perturbations do lead to a rejection of the overidentifying restrictions.

In the bottom panel of the table we estimate the model with the restriction $\beta = 1$. In every case, the restricted model is rejected by the overidentification test.[36]

[36] In columns 2 and 3, the exponential results are nearly identical to the benchmark case; the point estimates do not differ at all. This occurs because simulated exponential households accumulate zero illiquid wealth at the estimated $\hat{\delta}$ in both the benchmark case and in the perturbed cases.

5.3.2 The coefficient of relative risk aversion

We next examine the effect of varying the coefficient of relative risk aversion ρ. Economists disagree about how to calibrate ρ. In order to account for the equity premium puzzle, the consumption CAPM requires $\rho > 10$ (Kocherlakota 1996). But most lifecycle consumption models assume $\rho \in [0.5, 5]$, consistent with typical introspective choices about hypothetical large gambles (Mankiw and Zeldes 1991). Using a structural approach, Gourinchas and Parker (2002) identify ρ from lifecycle consumption profiles. For different specifications they find ρ between 0.1 and 5.3, with a benchmark estimate of 0.51. When liquidity constraints do not bind and preferences are time-separable, ρ is the inverse of the elasticity of intertemporal substitution. Euler Equation estimates of this elasticity range roughly between 0 and 1 (Hall 1988). Further complicating the picture, recent theoretical work casts doubt on the prevailing approach to modeling risk attitudes. Kahneman and Tversky (1979) and others propose and use models of loss aversion that imply first-order risk aversion. Rabin (2000) argues that seemingly-reasonable attitudes toward small gambles imply totally unreasonable attitudes toward larger gambles in an expected utility model with second-order risk aversion. Chetty (2004) proposes that consumption commitments could cause different local and global levels of risk aversion.

Recall that we adopted $\rho = 2$ for our benchmark. We now examine the effect of adopting $\rho = 1$ and $\rho = 3$. Column 1 of Table 4B reports the effect of assuming $\rho = 1$. We find that $\hat{\beta}$ and $\hat{\delta}$ both rise relative to their benchmark estimates. With less curvature in the utility function consumers are more willing to consume early in life and retire in relative poverty (and thereby more willing to borrow on their credit cards and less willing to accumulate assets). Raising β and δ offsets these effects. In this specification, $\hat{\beta}$ is only marginally significantly different from 1. Finally, assuming $\rho = 1$ generates a rejection of the overidentifying restrictions.

Column 3 of Table 4B reports the effect of assuming $\rho = 3$. As one would predict from the previous experiment, we now find that $\hat{\beta}$ and $\hat{\delta}$ both fall relative to their benchmark estimates. As a result the null hypothesis of $\beta = 1$ is rejected with a t-statistic of 3.2. Finally, assuming $\rho = 3$ does not lead to a rejection of the overidentifying restrictions. It is not clear what larger values of ρ would imply, though the gradient that we observe in our simulations suggests that raising the value of ρ lowers the estimated value of β.

Given the sensitivity of the model's quantitative results to the assumed value of ρ, we also estimate ρ simultaneously with β and δ. The results are reported in Column 4 of Table 4B. In this case, β, δ, and ρ are all precisely identified. We find $\hat{\rho} = 0.220$ with a standard error of 0.066. The model wants a low value of ρ, since the empirical moments reveal very little precautionary savings – i.e. savings in liquid form. Recall that credit card borrowing reflects a decumulation of liquid wealth.

At the estimated, low value of ρ, the short-run discount factor $\hat{\beta}$ rises, but $\hat{\beta}$ is still significantly below unity. Higher short-term impatience is still needed to reproduce both the wealth and credit card moments. When ρ is estimated, the standard errors on $\hat{\beta}$ and $\hat{\delta}$ fall. Goodness-of-fit improves substantially and the overidentification test is not rejected.

The lower half of Column 4 reports simultaneous estimates of ρ and δ, holding β fixed at 1. The resulting estimate of $\hat{\rho}$ is again very low. The estimate is imprecise, and risk neutrality is not rejected. To interpret this, recall that when (exponential) discounting is sufficient to match the credit card moments, lifecycle accumulation is minimal.

Similar results from optimally weighted estimation with ρ are reported in Column 5 of Table 4B.

5.3.3 Returns to Scale in Household Size

The household utility function depends on the number of people in the household and the intra-household returns to scale. In the benchmark case, we assume that the effective household size n_t at time t equals two spouses, plus the number of adult dependents, plus 0.4 times the number of child dependents.

In this subsection we analyze two modifications to these assumptions. First, we implement the "square root scale," (Ruggles 1990, Atkinson, Rainwater and Smeeding 1995), in which period utility equals $u(C_t, Z_t, N_t) = N_t \cdot \frac{\left(\frac{C_t + \gamma Z_t}{\sqrt{N_t}}\right)^{1-\rho} - 1}{1-\rho}$, where N_t is the total number of people in the household. This formulation treats household heads, child dependents, and adult dependents symmetrically, but still captures returns to scale since the square root function is concave. The effect of this change is reported in Column 6 of Table 4A. In the quasi-hyperbolic case, we find that β falls slightly, δ rises slightly, and goodness of fit worsens slightly. The overidentification test fails at the 5% level. In the exponential case, δ rises slightly.

Second, we study a "partial individual mortality" assumption. In the benchmark case mortality occurs at the level of the household, so the two household heads die simultaneously.[37] In contrast, under partial individual mortality we incorporate individual mortality directly into n_t.[38] We assume that utility depends continuously on the expected number of heads that remain alive until age 84, at which point the expected number of surviving heads falls below 1. After age 84, under partial individual mortality, we retain our original assumption that mortality is a discrete event. This change generates the estimates in Column 7 of Table 4A. Relative to the benchmark, β falls and δ rises. Now the overidentification restrictions are rejected.

5.3.4 Compound Cases

We also consider four "compound cases" to evaluate the consequences of letting γ, r^{CC}, ρ, and intrahousehold returns to scale reinforce each other. In Case A, we consider the effect of simultaneously assuming $\gamma = 3.38\%$, $r^{CC} = 13\%$, and $\rho = 3$; in Case B we assume $\gamma = 6.59\%$, $r^{CC} = 10\%$, and $\rho = 1$; in Case C we assume $\gamma = 3.38\%$, $r^{CC} = 13\%$, $\rho = 5$, the square root scale, and "individual mortality;" and in Case D we assume $\gamma = 6.59\%$, $r^{CC} = 10\%$, and $\rho = 0.2196$.

Case A combines three perturbations of first stage parameters that we discussed and reported above. In isolation, each of those perturbations lowers the estimates of β. In Column 6 of Table 4B we report that their combined effect results in estimates of $\hat{\beta} = 0.375$ and $\hat{\delta} = 0.972$. When we restrict β to equal 1 (Column 6) we estimate $\hat{\delta} = 0.770$. With such a low $\hat{\delta}$, the exponential discounting model predicts credit card borrowing and excess sensitivity quite well, but predicts approximately no wealth accumulation.

In Case B we combine the three perturbations that are opposite to those in Case A, and we estimate $\hat{\beta} = 0.908$ and $\hat{\delta} = 0.943$ (Column 7 of Table 4B). This estimate carries a small standard error, implying that even under aggressive assumptions about γ, r^{CC}, and ρ, we still find that $\hat{\beta}$ is significantly less than 1.

[37] This assumption induces two off-setting biases (Laibson et al. 2003). First, it rules out partial intra-household longevity insurance through independent mortality outcomes. Second, it does not account for the rise in per-capita consumption needs when one of the two household heads dies.

[38] For comparability to our benchmark case, we use our benchmark assumptions about returns to scale in the household, in contrast to the square root scale.

The third compound case, Case C, combines all the perturbations discussed above that have decreased $\hat{\beta}$, and also assumes $\rho = 5$. Now, reported in Column 8, we obtain $\hat{\beta} = 0.178$ and $\hat{\delta} = 0.997$. Goodness-of-fit worsens, and the overidentification test is rejected.

Finally, in Case D we adopt $\rho = 0.2196$–the value of ρ we obtained when estimating β, δ, and ρ simultaneously–and we also assume $\gamma = 6.59\%$ and $r^{CC} = 10\%$. As expected, this is the combination of assumptions that pushes β closest to 1. We find $\hat{\beta} = 0.947$ and $\hat{\delta} = 0.946$. However, the standard error on $\hat{\beta}$ has shrunk, as in Case B and in the simultaneous estimation of β, δ, and ρ reported above, so even in this case the $\beta = 1$ null hypothesis is rejected with a t-statistic of 3.31.

In the bottom half of Columns 6-9 we report estimates in the compound cases for the restricted exponential model and find that $\hat{\delta}$ ranges from 0.620 to 0.936. Case B is the most successful case for the exponential model among all of the simulations that we perform. Nevertheless, this simulation also rejects the overidentification restrictions.

6 Extensions

This paper's findings suggest several directions for future work.

6.1 Naivete

Strotz (1955), Akerlof (1991) and O'Donoghue and Rabin (1999a, 1999b) propose that decision makers with dynamically inconsistent preferences make current choices under the "naive" belief that later selves will act in the interests of the current self. Angeletos et al. (2001) find that naive and sophisticated quasi-hyperbolic agents behave similarly in consumption models like the one discussed in this paper.

However, two puzzles remain that perhaps a model of naivete could address. First, the sophisticated quasi-hyperbolics in these simulations would be better off if they had no access to credit cards. Specifically, according to a comparison of value functions, at age 20, sophisticated quasi-hyperbolics would be willing to pay $2000 to get rid of their credit cards immediately and *never* have access to them in the future. This begs the question of why only a tiny fraction of consumers cut up their credit cards.

Second, the spread between the cost of funds and the credit card interest rate is "too high."

As Ausubel (1991) has pointed out, the spread cannot be accounted for by standard explanations like default probabilities; instead consumers seem to pick their credit card under the naive belief that they will not borrow in the future. Teaser rates may also be explained by a model of naive quasi-hyperbolic discounting (Shui and Ausubel 2004).[39]

6.2 Further generalization of intertemporal preferences

Standard models study the one-parameter exponential discount function, and we have explored a two-parameter generalization. We may find that richer representations better describe discounting patterns. It is also possible that the nature of discounting changes over the course of the lifecycle,[40] providing another possible framework for explaining patterns of wealth accumulation and credit card borrowing.

6.3 Heterogeneity

Another important direction for future work is to relax the assumption of homogeneous preferences. Specifically, one might wonder whether two groups of exponential consumers, one patient and the other impatient, could account for the facts. To us, the data suggest that there is substantial heterogeneity in the population, but that the heterogeneity does not explain why the median household both borrows on its credit cards and invests in illiquid assets. Laibson et al. (2003) find that credit card borrowing is pervasive across the entire wealth distribution. Nevertheless, a model with heterogeneous preferences may resolve some of the empirical tensions discussed in this paper.

Interest rate heterogeneity might also explain the lifecycle wealth accumulation and credit card borrowing facts. A model of exponential consumers facing sharply different interest rates could simultaneously match the frequencies of credit card borrowing and the levels of midlife wealth accumulation. But actual interest rates paid on credit cards do not correlate with wealth holdings. Households surveyed in the 1998 Survey of Consumer Finances were asked about the interest rate they paid on their credit cards. Households in the top quartile of the distribution of wealth reported an average nominal rate of 13.27%, whereas households in the lowest quartile of wealth reported an

[39] See also the theoretical contribution of DellaVigna and Malmendier (2004).

[40] Our assumption that marginal utility varies with demographics (and thus with age) is a special case of the general class of age-specific discounting studied by Attanasio and Weber (1995) and Attanasio et al. (1999).

average of 13.59%.[41] Hence, households with substantial wealth do not face low credit card interest rates – e.g. teaser rates – that could explain their frequent credit card borrowing.

6.4 Institutional assumptions

The assets in the model are stylized and it would be natural to make our institutional assumptions more realistic. In a richer model households would be able to declare bankruptcy, default on their credit card debt, sell their Z assets, borrow against their Z assets, and engage in many other types of financial transactions. We are most interested in exploring changes in assumptions about Z, since the Z asset plays a central role in driving our results. For example, if the Z asset were perfectly liquid our model would be unable to explain credit card borrowing (since consumers with liquid retirement assets would spend down these liquid assets rather than borrowing on a high-interest credit card). Hence, the illiquidity of Z is critical for the predictive power of the model. Illiquidity of Z also mimics the optimal mechanisms that have recently been derived for quasi-hyperbolic consumers (Amador et al. 2004).

Other work has explored the consequences of making Z less illiquid (i.e., by introducing a transaction cost for liquidating Z).[42] This does not change the model's qualitative predictions. However, such transaction cost models may be too crude to capture the implications of collateralized borrowing against Z. For example, home equity lines of credit are rapidly gaining popularity in the U.S. and such instruments make heretofore illiquid home equity relatively easy to access. If home equity lines of credit provide immediate liquidity with little or no transactions costs, then Z is really not an illiquid asset. However, if applying for a home equity line of credit generates immediate effort costs as well as bureaucratic delays, then quasi-hyperbolic consumers will prefer to use their credit cards rather than applying for such credit lines (Laibson 1997).

To explore these issues, researchers should develop high frequency models in which the units of time are days and not years (and β multiplies all utility flows starting tomorrow). In such models the nature of the effort costs and the bureaucratic delays will play an important role. For example, if an application for a home equity credit line yields immediate effort costs and only delayed liquidity

[41] The means refer to high school educated households with outstanding balances on their credit cards.

[42] Laibson et al. (2003) report simulations in which liquidating part of the illiquid asset generates a fixed cost of $10,000 and a proportional cost of 10%.

(say in a week), such credit may *not* be tempting to quasi-hyperbolic consumers. However, if a home equity application has vanishingly small effort costs and yields immediate liquidity, then such a loan will be appealing to quasi-hyperbolic consumers and the model will then predict that splurges will *not* be funded from credit cards.

Such high frequency models are beyond the scope of the current paper. As home equity lines of credit and other technologies for making illiquid assets liquid become increasingly popular (only 7.8% of homeowners in the 1998 SCF had such loans), it will become more important for economists to study the micro-structure and micro-timing of the loan application process. Though the parameters of this paper's model are identified off data from a time period when home equity was more illiquid than today, if consumers sharply devalue rewards that are delayed by only a few days, models of the credit application process will need to reflect these high frequency preferences.

6.5 Consumption shocks

In addition to income uncertainty, consumers also face stochastic shocks to preferences and consumption needs (Amador et al. 2004). Expenses for car repairs and health care, for example, often come unexpectedly. In principle this additional volatility could generate higher levels of (illiquid) wealth accumulation and credit card borrowing by promoting saving after unusually good shocks and borrowing after unusually bad shocks. We examine this possibility qualitatively by proxying for consumption shocks with increased income uncertainty. Specifically, we re-estimate β and δ after doubling the variances of the parameters of our calibrated income process.[43] This yields results very similar to the benchmark run discussed above: the estimate of β, 0.62, is significantly less than 1; the specification test on the quasi-hyperbolic model is borderline; and the specification test on the exponential model, when β is restricted to equal 1, is rejected. Though these results support the intuition that volatility can not explain why the same time-consistent consumers would both accumulate illiquid assets and borrow on credit cards, subsequent analyses might formally calibrate and incorporate stochastic shocks to the marginal utility of consumption.

[43] We consider this a plausible upper bound, as it increases the conditional and unconditional standard deviations of income by about 50%. Hubbard, Skinner and Zeldes (1995) and Palumbo (1999) estimate dynamic processes of health expenditures, finding conditional and unconditional standard deviations equivalent to roughly one-third and one-fifth, respectively, of the conditional and unconditional standard deviations in income that we document in the PSID.

6.6 Second stage moment sets

As is well known from GMM theory, the choice of moments can be crucial to the outcome of the analysis. We chose to focus on the four moments discussed above: % *visa*, *mean visa*, *CY*, *wealth*. We could have matched other moments, including the wealth-to-income ratio at ages *other* than the pre-retirement age that we study; or the consumption-to-income ratio at various points in the lifecycle; or the proportion of homeowners; or the fraction of wealth that is illiquid; and so on. *If* our model is correctly specified, analyzing different moments should not change our estimated parameter values. Hence, analyzing different moments is a potential test of our model and a priority for future research.

Data on synthetic cohort consumption profiles are used by Gourinchas and Parker (2002) to estimate a long-term discount factor of 0.949 (0.015) for their high-school-graduate group. We believe that low-frequency lifecycle dynamics for average consumption are a good way to identify long-run discount rates. Wealth accumulation provides an alternative strategy to identify long-run discount rates. These different estimation strategies are likely to produce similar results, since the wealth-based strategy we adopt generates simulations that exhibit the hump-shaped consumption profiles used by Gourinchas and Parker (2002).[44] Nevertheless, it would be useful to integrate both sets of moments in a single estimation exercise. The calibrated quasi-hyperbolic model predicts a transitory consumption boom when credit cards are first acquired. One could empirically test this prediction.

Illiquidity of investments provides another source of identifying data (Angeletos et al 2001). Households invest very little of their wealth in liquid form. Even with an expansive definition of liquid assets, only 18.6% of total US household wealth is liquid.[45] Illiquidity may help to explain several puzzles involving consumption during retirement, including the measured pattern of anomalously slow decumulation of assets.

[44] See Angeletos et al. (2001) for simulations of lifecycle consumption.

[45] The methodology for calculating the share is analogous to that for calculating % *Visa*, *mean Visa*, and *wealth* described in Appendix A. The definition includes cash, checking and savings accounts, money market accounts, call accounts, CDs, bonds, stocks and mutual funds.

6.7 Policy experiments

Extensions of our structural model can be used to analyze public policies, including bankruptcy laws, credit card regulations, social insurance systems, and pensions policies (e.g., forced savings, automatic enrollment, annuitization, etc.). The positive and normative consequences of such policies can be analyzed by incorporating these policies into our structural model. With estimates in hand, out-of-sample predictions are possible.

7 Conclusion

This paper estimates time preferences using a structural model and field data. U.S. households accumulate large stocks of wealth before retirement, borrow actively on credit cards, and exhibit excess sensitivity of consumption to predictable movements in income. To explain these phenomena the MSM procedure estimates $\beta = 0.703$ and $\delta = 0.958$. Intuitively, the implied low long-term discount rate ($-\ln \delta = 4.3\%$) accounts for observed levels of (illiquid) wealth accumulation. The high short-term discount rate ($-\ln \beta\delta = 39.5\%$) explains the observed frequency and levels of credit card borrowing and excess sensitivity of consumption. Our benchmark specification fails to reject the overidentification restrictions. The MSM procedure does reject the restriction to exponential discounting ($\beta = 1$).

Our parameter estimates are sensitive to the calibration choices, and some calibrations lead to a failure of the overidentification tests. In addition, our economic environment is stylized and future work should enrich the realism of our modelling framework. However, the evidence reported here suggests that consumption-savings models will better match field data when the models incorporate discount rates in the short run that exceed discount rates at longer horizons. Structural estimation using field data is likely to be a useful complement to laboratory studies that measure time preferences.

A Second-Stage Moments Appendix

We now discuss the procedures we use to construct the second-stage moments. We use the SCF to derive $wealth$, $\%\ Visa$ and $mean\ Visa$, and the PSID to construct CY. Procedures are very similar for the share of liquid assets. All quantities are deflated to 1990 dollars.

A.1 SCF Moments

We use the 1983, 1989, 1992, 1995, and 1998 SCFs to compute the $wealth$ moment. We derive $\%\ Visa$ and $mean\ Visa$ from the 1995 and 1998 SCFs. We control for cohort effects, household demographics, and business cycle effects to make the characteristics of the population and the simulated data fully analogous. We assign to households in our simulations the mean empirical cohort, demographic, and business cycle effects. We adopt the following procedures.

For each variable of interest x we first use weighted least squares, applying the SCF population weights that match our sample selections, to estimate

$$x_i = FE_i + BCE_i + CE_i + AE_i + \xi_i \tag{8}$$

Here FE_i is a family size effect that consists of three variables, the number of heads, the number of children, and the number of dependent adults in the household. BCE_i is a business cycle effect proxied by the unemployment rate in the household's region of residence. In 1983, the unemployment rate is the rate in the state of residence. In 1992, 1995, and 1998, it is the rate in the Census Division. In 1989 the nationwide rate was used because information on household location is not available in the public use data set. CE_i is a cohort effect that consists of a full set of five-year cohort dummies, AE_i is an age effect that consists of a full set of age dummies, and ξ_i is an error term.[46]

Next, we define the "typical" household to be identical to the simulated household (i.e. with head and spouse, exogenous age-varying numbers of children and adult dependents, an average cohort effect, and an average unemployment effect[47]). Then for each variable we create a new

[46] Following Gourinchas and Parker (2002) we attribute time effects to fluctuations in unemployment, but this approach–like any approach to separating age, cohort, and time effects–requires problematic identification assumptions. Deaton (1997) discusses other imperfect alternatives.

[47] These averages are the means used in the calibration of the income process, which is based on the PSID. Refer

variable $\widehat{x_i}$ that captures what x_i would have been had household i been typical. For example, if household i is identical to the "typical" household except for having more children, we set $\widehat{x_i} = x_i + \beta(\overline{nkids} - nkids_i)$, where β is the coefficient for number of kids in the regression above and $\overline{nkids}$ is the average number of children in a household as a function of the head's age. All moments were estimated using $\widehat{x_i}$.

For *wealth*, we restrict the sample to households with heads between ages 50-59. We include all real and financial wealth (e.g., home equity and CDs) as well as all claims on defined contribution pension plans (e.g., 401(k)). The measure does not include Social Security wealth and claims on defined benefit pension plans, since these flows appear in our estimated income process. If a household had a negative net balance in any illiquid asset, we set the balance equal to zero (e.g., we set home equity equal to the max of 0 and the value of home minus outstanding mortgages minus used portion of home equity lines of credit). Since there is no separate information on the amount borrowed against home equity lines of credit in the 1983 SCF, we assume that in that year no household had an outstanding home equity line balance.[48]

Let $\kappa = 10 \cdot \frac{2}{\pi}$. Then *wealth* is the mean of $\kappa \cdot \arctan\left(\frac{\widehat{x_i}}{\kappa}\right)$ in the sample, applying the SCF population weights. We use this arctan scaling in order to downweight outliers. This function has noteworthy properties. First, it is symmetric around the origin. Second, it is approximately linear in a neighborhood of the origin. Third, as $\widehat{x_i}$ gets very large, it asymptotes to 10. We compute the standard error of *wealth* directly from the sample values of $\kappa \cdot \arctan\left(\frac{\widehat{x_i}}{\kappa}\right)$.

To construct $\%\ Visa$ we create a dummy variable *hasdebt* equal to one for household i if i has a positive outstanding credit card balance in the SCF. We correct *hasdebt* to generate $\widehat{x_i}$. We then regress $\widehat{x_i}$ on a full set of age dummies. $\%\ Visa$ is a linear combination of the estimated coefficients on the age dummies, where the weights are derived from the same conditional survival probabilities we use in the simulations. The standard error is computed directly from the weights and the standard errors on the age dummy estimates.

Construction of $mean\ Visa$ is complicated by the fact that aggregate credit card borrowing data from the Fed indicate that 1995 and 1998 SCF borrowing magnitudes are biased downward

to Table 3 and Laibson et al. (2003) for details.

[48] In the 1983 SCF, 1.7% of homeowners with a high school degree reported having a credit line secured by home equity.

by a factor of three. We correct for this bias as follows. First we compute average outstanding interest-bearing balances. According to the Fed, aggregate debt outstanding at year-end 1995 and 1998 were $443 billion and $561 billion, respectively. From these figures we subtract an upper bound on the float (the balances that are still in their one-month grace period, which do not accrue interest). This upper bound is obtained by dividing total purchase volume, approximately $1 trillion in 1998, by 12. We then divide by the number of U.S. households with credit cards, using Census Bureau data on total households and SCF data on the percentage of households with cards. We obtain average household borrowing conditional on having a card of $5115 in 1995 and $6411 in 1998. These figures are consistent with those from a proprietary account-level data set analyzed by Gross and Souleles (2002a, 2002b).

In our simulations we focus on households headed by people with high school degrees, so next we use the SCF data on borrowing to scale the Fed average borrowing figure for just the high school educated group. In particular, we define α such that

$$debt_{all}^{Fed} = \alpha \cdot (w_{nhs} debt_{nhs}^{SCF} + w_{hs} debt_{hs}^{SCF} + w_{coll} debt_{coll}^{SCF})$$

with weights w_{nhs}, w_{hs}, and w_{coll} defined by the proportion of educational categories in the population (0.25, 0.5, 0.25, respectively) and $debt_{educ}^{source}$ equal to the average debt reported by *source* for educational group *educ*. Focusing now exclusively on the HS educational group, let $debt_i^{SCF}$ be the level of credit card debt reported in the SCF for household i. Let $debt_i = \alpha \cdot debt_i^{SCF}$ be the corrected credit card debt. Calculate age specific income means ($\overline{y}_t$) and create $debtinc_i$ as $debt_i/\overline{y}$ [49]. Then, we correct $debtinc_i$, creating $\widehat{x}_i$, and regress $\widehat{x}_i$ on a full set of age dummies. The moment $mean\ Visa$ is a linear combination of the estimated coefficients on the age dummies, again using the weights derived from the conditional survival probabilities used in the simulations. Again, the standard error is computed directly from the weights and the standard errors on the age dummy estimates.

Covariances between the SCF moments were constructed by jointly estimating the above means.

[49] When calculating the age-specific income means we group together ages 20-21, 70-74, 75-79, and 80 and over because we have very few observations at those ages.

A.2 PSID Moment

We use PSID data from 1978 to 1992 to estimate the CY moment. In the data, we define consumption to include food, rent, and utilities (the most general definition available in the PSID). The rental value of an owner-occupied home is assumed to be 5% of the value of the home. If the household neither owns nor rents, rent is the self-reported rental value of the home if it were rented.

We construct the CY moment by using 2SLS to estimate

$$\Delta \ln(C_{it}) = \alpha E_{t-1} \Delta \ln(Y_{it}) + X_{it}\beta + \varepsilon_{it},$$

where C_{it} is just food, rent, and utilities.[50] We assume an MA(1) process for the error term and instrument for $E_{t-1}\Delta \ln(Y_{it})$ with $\ln Y_{it-3}$ and $\ln Y_{it-4}$. The instruments are jointly sgnificant in the first stage regression, and the overidentification test does not reject the specification. The vector X_{it} includes age, cohort, and business cycle effects, the change in effective family size, the mortality rate, and lagged wealth.[51] Since wealth is observed in the PSID only in 1984 and 1989, in the other years we estimate wealth using the intertemporal budget constraint and a projected value of total consumption. Total consumption was projected from the PSID's partial measure using the CEX: in the CEX we regress total consumption on food, rent and utilities consumption, and then we use the coefficients to infer total consumption from the available PSID measure.

B MSM Procedure Appendix

Since $m(\theta; \widehat{\chi})$ is difficult to evaluate we replace it with an unbiased simulator, calculated by first taking J_s draws of the initial distribution and then constructing the corresponding simulated ex-

[50] Unfortunately, the PSID contains limited information on consumption. Our CY measure may be sensitive to the definition of consumption. Using the Consumer Expenditure Survey, Dynarski and Gruber (1997) report a lower sensitivity of food and housing expenditures than of total consumption to changes in the earnings of the head of the household. Hence, the comovement moment that we use in our analysis is biased down, which implies that our estimate of β is probably (slightly) biased up, since low β values generate more comovement. We conjecture that this bias in β is slight because the comovement moment has a minimal impact on our estimates due to the high standard errors that are associated with that moment.

[51] The wealth variable turned out to be significant at a 6% level. Its exclusion does not affect the estimate of CY, which slightly rises from 0.2311 to 0.2314.

pectations. Define $m_{J_s}(\theta;\widehat{\chi})$ as the vector of simulated moments. Now we can find the vector $\widehat{\theta}$ that minimizes $g'_{J_s}(\theta;\widehat{\chi})\,W g_{J_s}(\theta;\widehat{\chi})$, where $g_{J_s}(\theta;\widehat{\chi}) = \bar{m}_{J_m} - m_{J_s}(\theta;\widehat{\chi})$.

The first order condition for the second stage (incorporating the use of simulation) is given by

$$g'_{J_s\theta}\left(\widehat{\theta};\widehat{\chi}\right) W g_{J_s}\left(\widehat{\theta};\widehat{\chi}\right) = 0.$$

where $g_{J_s\theta}\left(\widehat{\theta};\widehat{\chi}\right) = \partial g_{J_s}\left(\widehat{\theta};\widehat{\chi}\right)/\partial\theta'$.

Following Gourinchas and Parker (2002) and Newey and McFadden (1994), an expansion of $g_{J_s}\left(\widehat{\theta};\widehat{\chi}\right)$ around θ_0 to first order leads to

$$g'_{J_s\theta}\left(\widehat{\theta};\widehat{\chi}\right) W\left[g_{J_s}(\theta_0;\widehat{\chi}) + g_{J_s\theta}(\theta_0;\widehat{\chi})\left(\widehat{\theta}-\theta_0\right)\right] = 0.$$

Rearranging terms and defining $\hat{J}_m$ as the (scalar) rate of convergence of $\widehat{\theta}$,

$$\sqrt{\hat{J}_m}\left(\widehat{\theta}-\theta_0\right) = -\left[g'_{J_s\theta}\left(\widehat{\theta};\widehat{\chi}\right) W g_{J_s\theta}(\theta_0;\widehat{\chi})\right]^{-1} g'_{J_s\theta}\left(\widehat{\theta};\widehat{\chi}\right) W\sqrt{\hat{J}_m}\,g_{J_s}(\theta_0;\widehat{\chi}).$$

Let $\Pi \equiv \left[g'_{J_s\theta}\left(\widehat{\theta};\widehat{\chi}\right) W g_{J_s\theta}(\theta_0;\widehat{\chi})\right]^{-1} g'_{J_s\theta}\left(\widehat{\theta};\widehat{\chi}\right) W$. Expanding $g_{J_s}(\theta_0;\widehat{\chi})$ around χ_0,

$$\sqrt{\hat{J}_m}\left(\widehat{\theta}-\theta_0\right) = -\Pi\left[\sqrt{\hat{J}_m}\,g_{J_s}(\theta_0;\chi_0) + \sqrt{\hat{J}_m}\,g_{J_s\chi}(\theta_0;\chi_0)\left(\widehat{\chi}-\chi_0\right)\right]. \qquad (9)$$

To evaluate Equation 9, first note that

$$\begin{aligned}\sqrt{\hat{J}_m}\,g_{J_s}(\theta_0;\chi_0) &= \sqrt{\hat{J}_m}\left[\bar{m}_{J_m} - m_{J_s}(\theta_0;\chi_0)\right]\\ &= \sqrt{\hat{J}_m}\left[\bar{m}_{J_m} - m(\theta_0;\chi_0)\right] + \sqrt{\hat{J}_m}\left[m(\theta_0;\chi_0) - m_{J_s}(\theta_0;\chi_0)\right]\end{aligned}$$

The two bracketed terms represent independent sets of draws from the same population. The first term equals $\sqrt{\hat{J}_m}\,g(\theta_0;\chi_0)$, which is asymptotically normally distributed: $\sqrt{\hat{J}_m}\,g(\theta_0;\chi_0) \to N(0,\Sigma_g)$. We estimate $\Omega_g = \frac{\Sigma_g}{J_m} = E\left[g(\theta_0;\chi_0)\,g(\theta_0;\chi_0)'\right]$ directly from its sample counterpart.[52] The second term represents the simulation error. At the true value of θ, the simulated moments were generated from a finite number of random draws from the true population. Therefore, the

[52] In fact, the (a,b)'th cell is $\Omega_g(a,b) = \frac{\Sigma_g(a,b)}{\min(J_{ma},J_{mb})}$.

second term is also asymptotically normal (as the size of the simulated sample goes to infinity) with mean 0 and variance $\hat{J}_m \frac{\Sigma_g}{J_s}$. Finally, since variation in the simulation and the data are independent, $\sqrt{\hat{J}_m} g_{J_s}(\theta_0; \chi_0) \to N\left(0, \left(1 + \frac{\hat{J}_m}{J_s}\right)\Sigma_g\right)$. To operationalize this expression for the variance, given the different numbers of observations J_m in the sample, we conservatively use the pairwise maximum numbers of observations, $\max(J_{ma}, J_{mb})$, to weight the (a,b)'th cell of Σ_g in the simulation correction.

Now turn to the second term of Equation 9. In the main text we have defined the variance of the first stage parameter estimates $\widehat{\chi}$ as $\Omega_\chi = E\left[(\widehat{\chi} - \chi_0)(\widehat{\chi} - \chi_0)'\right]$.

Thus, $\sqrt{\hat{J}_m} g_{J_s\chi}(\theta_0; \chi_0)(\widehat{\chi} - \chi_0) \to N\left(0, \hat{J}_m G_\chi \Omega_\chi G_\chi'\right)$, and $\sqrt{\hat{J}_m}\left(\widehat{\theta} - \theta_0\right) \to N(0, \Sigma_\theta)$, where Equation 9 implies

$$\Sigma_\theta = \left(G_\theta' W G_\theta\right)^{-1} G_\theta' W \left[\left(1 + \frac{\hat{J}_m}{J_s}\right)\Sigma_g + \hat{J}_m \cdot G_\chi \Omega_\chi G_\chi'\right] W G_\theta \left(G_\theta' W G_\theta\right)^{-1}, \tag{10}$$

by the asymptotic normality of $\widehat{\chi}$ and $g(\cdot)$ and by the Slutsky theorem, assuming zero covariance between the first and second stage moments. Dividing by $\hat{J}_m$ we obtain our key equation,

$$\Omega_\theta = Var\left(\hat{\theta}\right) = \left(G_\theta' W G_\theta\right)^{-1} G_\theta' W \left[\Omega_g + \Omega_g^s + G_\chi \Omega_\chi G_\chi'\right] W G_\theta \left(G_\theta' W G_\theta\right)^{-1}.$$

Standard errors reported in the text and tables equal the square roots of the diagonal elements of Ω_θ.

References

Ainslie, George, Picoeconomics, Cambridge: Cambridge University Press, 1992.

Akerlof, George, "Procrastination and Obedience," American Economic Review, May 1991, 81 (2), 1–19.

Altonji, Joseph and Lewis Segal, "Small-Sample Bias in GMM Estimation of Covariance Structures," Journal of Business and Economic Statistics, 1996, 14 (3), 353–66.

Amador, Manuel, Ivan Werning, and George-Marios Angeletos, "Commitment Vs. Flexibility," 2004. MIT mimeo.

Angeletos, George-Marios, David Laibson, Andrea Repetto, Jeremy Tobacman, and Stephen Weinberg, "The Hyperbolic Consumption Model: Calibration, Simulation, and Empirical Evaluation," Journal of Economic Perspectives, Summer 2001, 15 (3), 47–68.

Atkinson, Anthony, Lee Rainwater, and Timothy Smeeding, "Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study (LIS)," Technical Report Social Policy Studies 18, OECD 1995.

Attanasio, Orazio and Guglielmo Weber, "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey," Journal of Political Economy, 1995, 103 (6), 1121–1157.

——, James Banks, Costas Meghir, and Guglielmo Weber, "Humps and Bumps in Lifetime Consumption," Journal of Business and Economic Statistics, January 1999, 17 (1), 22–35.

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

Bernheim, B. Douglas, Jonathan Skinner, and Steven Weinberg, "What Accounts for the Variation in Retirement Saving Among U.S. Households?," American Economic Review, September 2001, 91 (4), 832–857.

Bertaut, Carol and Michael Haliassos, "Debt Revolvers for Self Control," 2001. University of Cyprus mimeo.

Browning, Martin and Annamaria Lusardi, "Household Saving: Micro Theories and Micro Facts," Journal of Economic Literature, December 1996, 34 (4), 1797–1855.

Carroll, Christopher, "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," Brookings Papers on Economic Activity, 1992, 1992 (2), 61–156.

——, "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," Quarterly Journal of Economics, February 1997, 112 (1), 1–55.

—— and Andrew Samwick, "The Nature of Precautionary Wealth," Journal of Monetary Economics, 1997, 40, 41–71.

—— and Lawrence Summers, "Consumption Growth Parallels Income Growth: Some New Evidence," in B. Douglas Bernheim and John B. Shoven, eds., National Saving and Economic Performance, Chicago: Chicago University Press, 1991.

Chetty, Raj, "Consumption Commitments, Unemployment Durations, and Local Risk Aversion," 2004. NBER Working Paper 10211.

Chung, Shin-Ho and Richard Herrnstein, "Choice and Delay of Reinforcement," Journal of the Experimental Analysis of Behavior, 1967, 10, 67–74.

Deaton, Angus, "Saving and Liquidity Constraints," Econometrica, 1991, 59 (5), 1221–1248.

——, Understanding Consumption, Oxford: Oxford University Press, 1992.

——, The Analysis of Household Surveys: A Microeconometric Approach to Development Policy, Baltimore: Johns Hopkins University Press, 1997.

DellaVigna, Stefano and Ulrike Malmendier, "Contract Design and Self-Control: Theory and Evidence," Quarterly Journal of Economics, May 2004, 119 (2), 353–402. Harvard University mimeo.

Duffie, Darrell and Ken Singleton, "Simulated Moments Estimation of Markov Models of Asset Prices," Econometrica, 1993, 61 (4), 929–952.

Dynarski, Susan and John Gruber, "Can Families Smooth Variable Earnings?," Brookings Papers on Economic Activity, 1997, pp. 229–305.

Engen, Eric, William Gale, and Cori Uccello, "The Adequacy of Household Saving," Brookings Papers on Economic Activity, 1999, 1999 (2), 65–165.

——, ——, and John Karl Scholz, "Do Saving Incentives Work?," Brookings Papers on Economic Activity, 1994, 1994 (1), 85–180.

Fang, Hanming and Dan Silverman, "Time Inconsistency and Welfare Program Participation: Evidence from the NLSY," 2004. Cowles Foundation Discussion Paper 1465, Yale University.

Flavin, Marjorie and Takashi Yamashita, "Owner-Occupied Housing and the Composition of the Household Portfolio," American Economic Review, March 2002, 92 (1), 345–362.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue, "Time Discounting and Time Preference: A Critical Review," Journal of Economic Literature, 2002, 40 (2), 351–401.

French, Eric, "The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behavior," Review of Economic Studies, 2005, 72 (2), 395–437.

Gourinchas, Pierre-Olivier and Jonathan Parker, "Consumption over the Life Cycle," Econometrica, January 2002, 70 (1), 47–89.

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

—— and ——, "An Empirical Analysis of Personal Bankruptcy and Delinquency," Review of Financial Studies, Spring 2002, 15 (1), 319–347.

Hall, Robert, "Intertemporal Substitution in Consumption," Journal of Political Economy, 1988, 96 (2), 339–357.

Hausman, Jerry, "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables," Bell Journal of Economics, 1979, 10 (1), 33–54.

Hubbard, Glenn, Jonathan Skinner, and Stephen Zeldes, "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving," Carnegie-Rochester Conference Series on Public Policy, June 1994, 40, 59–126.

——, ——, and ——, "Precautionary Saving and Social Insurance," Journal of Political Economy, 1995, 103 (2), 360–399.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," Econometrica, March 1979, 47 (2), 263–291.

Kocherlakota, Narayana, "The Equity Premium: It's Still a Puzzle," Journal of Economic Literature, March 1996, 34 (1), 42–71.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," Quarterly Journal of Economics, 1997, 62 (2), 443–478.

——, Andrea Repetto, and Jeremy Tobacman, "Self-Control and Saving for Retirement," Brookings Papers on Economic Activity, 1998, (1), 91–196.

——, ——, and ——, "A Debt Puzzle," in Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford, eds., Knowledge, Information, and Expectations in Modern Economics: In Honor of Edmund S. Phelps, Princeton: Princeton University Press, 2003.

Lawrence, Emily, "Poverty and the Rate of Time Preference: Evidence from Panel Data," Journal of Political Economy, 1991, 99 (1), 54–77.

Mankiw, N. Gregory and Stephen P. Zeldes, "The Consumption of Stockholders and Non-Stockholders," Journal of Financial Economics, 1991, 29 (1), 97–112.

McFadden, Daniel, "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," Econometrica, 1989, 57 (5), 995–1026.

Newey, Whitney and Daniel McFadden, "Large Sample Estimation and Hypothesis Testing," in Robert Engle and Daniel McFadden, eds., Handbook of Econometrics, Volume IV, Elsevier Science, 1994, pp. 2111–2245.

O'Donoghue, Ted and Matthew Rabin, "Doing It Now or Later," American Economic Review, March 1999, 89 (1), 103–124.

____ and ____ , "Incentives for Procrastinators," Quarterly Journal of Economics, August 1999, 114 (3), 769–816.

Pakes, Ariel and David Pollard, "Simulation and the Asymptotics of Optimization Estimators," Econometrica, 1989, 57 (5), 1027–1057.

Palumbo, Michael, "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle," Review of Economic Studies, April 1999, 66 (2), 395–421.

Parker, Jonathan, "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes," American Economic Review, September 1999, 89 (4), 959–973.

Paserman, Daniele, "Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation," 2004. CEPR Discussion Paper 4396.

Phelps, Edmund and Robert Pollak, "On Second-Best National Saving and Game-Equilibrium Growth," Review of Economic Studies, 1968, 35, 185–199.

Rabin, Matthew, "Risk Aversion and Expected Utility Theory: A Calibration Theorem," Econometrica, 2000, 68 (5), 1281–1292.

Ruggles, Patricia, Drawing the Line: Alternative Poverty Measures and Their Implications for Public Policy, Washington, DC: The Urban Institute Press, 1990.

Samwick, Andrew A., "Discount Rate Heterogeneity and Social Security Reform," Journal of Development Economics, October 1998, 57 (1), 117–146.

Shea, John, "Union Contracts and the Life-Cycle/Permanent Income Hypothesis," American Economic Review, March 1995, 85 (1), 186–200.

Shui, Haiyan and Lawrence M. Ausubel, "Time Inconsistency in the Credit Card Market," Econometric Society North American Summer Meetings 176 May 3 2004.

Stern, Steven, "Simulation-Based Estimation," Journal of Economic Literature, December 1997, 35 (4), 2006–2039.

Strotz, Robert, "Myopia and Inconsistency in Dynamic Utility Maximization," Review of Economic Studies, 1955, 23 (3), 165–180.

U.S. Census Bureau, Statistical Abstract of the United States 1995.

U.S. National Center for Health Statistics, Vital Statistics of the United States, Department of Health and Human Services, 1994.

Viscusi, Kip and Michael Moore, "Rates of Time Preference and Valuation of the Duration of Life," Journal of Public Economics, 1989, 38 (3), 297–317.

Warner, John and Saul Pleeter, "The Personal Discount Rate: Evidence from Military Downsizing Programs," American Economic Review, 2001, 91 (1), 33–53.

West, Kenneth, Ka-Fu Wong, and Stanislav Anatolyev, "Instrumental Variables Estimation of Heteroskedastic Linear Models Using All Lags of Instruments," 2004. University of Wisconsin mimeo.

Zeldes, Stephen, "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," Quarterly Journal of Economics, 1989, 104 (2), 275–298.

TABLE 1
SECOND-STAGE MOMENTS

Description and Name	$\overline{m}_{J_m}$	se($\overline{m}_{J_m}$)
% Borrowing on Visa: *"% Visa"*	0.678	0.015
Mean (Borrowing$_t$ / mean(Income$_t$)): *"mean Visa"*	0.117	0.009
Consumption-Income Comovement: *"CY"*	0.231	0.112
Average weighted $\frac{wealth}{income}$: *"wealth"*	2.60	0.13

Source: Authors' calculations based on data from the Survey of Consumer Finances, the Federal Reserve, and the Panel Study on Income Dynamics. Calculations pertain to households with heads who have high school diplomas but not college degrees. The variables are defined as follows: *% Visa* is the fraction of U.S. households borrowing and paying interest on credit cards (SCF 1995 and 1998); *mean Visa* is the average amount of credit card debt as a fraction of the mean income for the age group (SCF 1995 and 1998, weighted by Fed aggregates); *CY* is the marginal propensity to consume out of anticipated changes in income (PSID 1978-92); and *wealth* is the weighted average wealth-to-income ratio for households with heads aged 50-59 (SCF 1983-1998).

TABLE 2

FIRST STAGE ESTIMATION RESULTS

Demographics

Number of children

$k = \beta_0 * \exp(\beta_1 * age - \beta_2 * (age^2)/100) + \varepsilon$

β_0	β_1	β_2
0.006	0.324	0.005
(0.001)	(0.005)	(0.007)

Number of dependent adults

$a = \beta_0 * \exp(\beta_1 * age - \beta_2 * (age^2)/100) + \varepsilon$

β_0	β_1	β_2
8.0e-9	0.727	0.007
(0.000)	(0.016)	(0.016)

Liquid assets and noncollateralized debt

Credit limit λ

0.318
(0.017)

Return on positive liquid assets R

1.0279
(0.024)

Credit card interest rate R^{cc}

1.1152
(0.009)

Illiquid Assets

Consumption flow as a fraction of assets γ

0.05
-

Preference Parameter

Coefficient of relative risk aversion ρ

2
-

Income from transfers and wages

Income process - In the labor force

$y = \ln(Y) = \beta_0 + \beta_1 * age + \beta_2 * (age^2/100) + \beta_3 * (age^3/10000) + \beta_4 * Nheads + \beta_5 * Nchildren + \beta_6 * Ndep.adults + \xi^W$

$\xi^W_t = \eta_t + \upsilon_t = \alpha\eta_{t-1} + \varepsilon_t + \upsilon_t$

β_0	β_1	β_2	β_3	β_4	β_5	β_6	α	σ^2_ε	σ^2_υ
7.439	0.118	-0.201	0.081	0.548	-0.033	0.170	0.782	0.029	0.026
(0.340)	(0.021)	(0.050)	(0.035)	(0.019)	(0.005)	(0.008)	(0.017)	(0.008)	(0.011)

Income Process - Retired

$y = \ln(Y) = \beta_0 + \beta_1 * age + \beta_2 * Nheads + \beta_3 * Nchildren + \beta_4 * Ndep.adults + \xi^R$

β_0	β_1	β_2	β_3	β_4	$\sigma^2_{\xi^R}$
8.433	-0.002	0.554	0.199	0.204	0.051
(0.849)	(0.013)	(0.084)	(0.172)	(0.102)	(0.013)

Retirement age T

63
(0.730)

Source: Authors' estimation, exactly following Laibson, Repetto, and Tobacman (2003), based on data from the PSID, SCF, FRB, and American Bankruptcy Institute, for households with heads who have high school diplomas but not college degrees.

Note: Standard errors in parentheses. The constant of the deterministic component of income includes a year of birth cohort effect and a business cycle effect proxied by the unemployment rate.

The dynamics of income estimation includes a household fixed effect.

Illiquid asset consumption flows and the coefficient of relative risk aversion are assumed to be exactly known in the context of the first stage. We examine sensitivity to these parameters in Subsection 5.3 on Robustness.

This table only reports standard errors, but the full covariance matrix is used in the second-stage estimation.

TABLE 3

BENCHMARK STRUCTURAL ESTIMATION RESULTS

	(1) Hyperbolic	(2) Exponential	(3) Hyperbolic Optimal Wts	(4) Exponential Optimal Wts	(5) Data
Parameter estimates $\hat{\theta}$					
$\hat{\beta}$	0.7031	1.0000	0.7150	1.0000	-
s.e. (i)	(0.1093)	-	(0.0948)	-	-
s.e. (ii)	(0.1090)	-	-	-	-
s.e. (iii)	(0.0170)	-	-	-	-
s.e. (iv)	(0.0150)	-	-	-	-
$\hat{\delta}$	0.9580	0.8459	0.9603	0.9419	-
s.e. (i)	(0.0068)	(0.0249)	(0.0081)	(0.0132)	-
s.e. (ii)	(0.0068)	(0.0247)	-	-	-
s.e. (iii)	(0.0010)	(0.0062)	-	-	-
s.e. (iv)	(0.0009)	(0.0056)	-	-	-
Second-stage moments					
% Visa	0.634	0.669	0.613	0.284	0.678
mean Visa	0.167	0.150	0.159	0.049	0.117
CY	0.314	0.293	0.269	0.074	0.231
wealth	2.69	-0.05	3.22	2.81	2.60
Goodness-of-fit					
$q(\hat{\theta},\hat{\chi})$	67.2	436	2.48	34.4	-
$\xi(\hat{\theta},\hat{\chi})$	3.01	217	8.91	258.7	-
p-value	0.222	<1e-10	0.0116	<2e-7	-

Source: Authors' calculations.

Note on standard errors: (i) includes both the first stage correction and the simulation correction, (ii) includes just the first stage correction, (iii) includes just the simulation correction, and (iv) includes neither correction.

TABLE 4A

ROBUSTNESS: RATES OF RETURN AND RETURNS TO SCALE

	(1) Benchmark	(2) $\gamma = 3.38\%$	(3) $\gamma = 6.59\%$	(4) $r^{CC} = 10\%$	(5) $r^{CC} = 13\%$	(6) Square Root Scale	(7) Partial Individual Mortality
Hyperbolic							
Parameter Estimates $\hat{\theta}$							
$\hat{\beta}$	0.7031	0.5071	0.8024	0.7235	0.6732	0.6708	0.6232
s.e. (i)	(0.1093)	(0.4410)	(0.0614)	(0.1053)	(0.1167)	(0.1530)	(0.1219)
$\hat{\delta}$	0.9580	0.9731	0.9425	0.9567	0.9595	0.9626	0.9651
s.e. (i)	(0.0068)	(0.0188)	(0.0093)	(0.0071)	(0.0045)	(0.0086)	(0.0069)
Goodness-of-fit							
$q(\hat{\theta},\hat{\chi})$	67.2	108.4	49.7	64.1	70.7	72.1	76.6
$\xi(\hat{\theta},\hat{\chi})$	3.01	16.79	5.27	12.09	10.97	9.05	18.48
p-value	0.222	0.0002	0.0717	0.0024	0.0041	0.0108	0.0001
Exponential							
Parameter Estimates $\hat{\theta}$							
$\hat{\beta}$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
s.e. (i)	-	-	-	-	-	-	-
$\hat{\delta}$	0.8459	0.8459	0.8459	0.8520	0.8354	0.8499	0.8442
s.e. (i)	(0.0249)	(0.0249)	(0.0250)	(0.0267)	(0.0262)	(0.0308)	(0.0407)
Goodness-of-fit							
$q(\hat{\theta},\hat{\chi})$	435.6	435.6	435.6	434.7	436.6	436.8	443.6
$\xi(\hat{\theta},\hat{\chi})$	217	217	263	177	339	201	374
p-value	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10

Source: Authors' calculations.

Note: The benchmark assumes γ=5%, r^{CC}=11.52%, and ρ=2. Columns (3) through (7) make perturbations one at a time. The square root scale assumes that, if the household size is *n*, period utility equals *n*u*(*c/sqrt*(*n*)), where *c* is total consumption. "Partial individual mortality" puts mortality into *n*, so that *n* falls in proportion to unconditional survival probabilities, until on average only one spouse remains in the household.

TABLE 4B

ROBUSTNESS: RISK AVERSION AND COMPOUND CASES

	(1) $\rho = 1$	(2) $\rho = 2$ (benchmark)	(3) $\rho = 3$	(4) Estimate ρ	(5) Estimate ρ, optimal wts	(6) Compound Case A	(7) Compound Case B	(8) Compound Case C	(9) Compound Case D
Hyperbolic									
Parameter Estimates $\hat{\theta}$									
$\hat{\beta}$	0.8186	0.7031	0.5776	0.8983	0.8989	0.3750	0.9075	0.1779	0.9473
s.e. (i)	(0.0959)	(0.1093)	(0.1339)	(0.0184)	(0.0211)	(0.4859)	(0.0285)	(0.2298)	(0.0159)
$\hat{\delta}$	0.9610	0.9580	0.9545	0.9616	0.9625	0.9717	0.9434	0.9968	0.9457
s.e. (i)	(0.0037)	(0.0068)	(0.0096)	(0.0025)	(0.0039)	(0.0228)	(0.0059)	(0.0145)	(0.0031)
$\hat{\rho}$	1.0000	2.0000	3.0000	0.2196	0.2299	3.0000	1.0000	5.0000	0.2196
s.e. (i)	-	-	-	(0.0655)	(0.0706)	-	-	-	-
Goodness-of-fit									
$q(\hat{\theta}, \hat{\chi})$	63.0	67.2	67.7	6.90	2.77	106.1	38.9	112.5	7.66
$\xi(\hat{\theta}, \hat{\chi})$	7.97	3.01	1.85	0.91	2.77	16.1	7.52	16.68	1.62
p-value	0.019	0.222	0.397	0.341	0.096	0.0003	0.023	0.0002	0.445
Exponential									
Parameter Estimates $\hat{\theta}$									
$\hat{\beta}$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
s.e. (i)	-	-	-	-	-	-	-	-	-
$\hat{\delta}$	0.8924	0.8459	0.7841	0.9102	0.9122	0.7695	0.9359	0.6200	0.9214
s.e. (i)	(0.0204)	(0.0249)	(0.0357)	(0.0078)	(0.0086)	(0.0262)	(0.0071)	(0.0560)	(0.0075)
$\hat{\rho}$	1.0000	2.0000	3.0000	0.2685	0.2825	3.0000	1.0000	5.0000	0.2685
s.e. (i)	-	-	-	(0.2419)	(0.1885)	-	-	-	-
Goodness-of-fit									
$q(\hat{\theta}, \hat{\chi})$	438.1	435.6	435.5	432.27	381.73	436.1	145.2	434.2	433.6
$\xi(\hat{\theta}, \hat{\chi})$	349	217	310	369.72	381.73	319.5	19.68	320.3	373.3
p-value	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10	0.0002	<1e-10	<1e-10

Source: Authors' calculations. Note: Columns (1) and (3) perturb only ρ. Columns (4) and (5) estimate ρ. Compound Case A assumes γ=3.38%, r^{CC}=13%, and ρ=3. Case B assumes γ=6.59%, r^{CC}=10%, and ρ=1. Case C assumes γ=3.38%, r^{CC}=13%, ρ=5, the square root scale, and "partial individual mortality." Case D for the exponential model assumes γ=6.59%, r^{CC}=10%, and ρ=0.2685 and Case D for the hyperbolic model assumes γ=6.59%, r^{CC}=10%, and ρ=0.2196. In the benchmark, γ=5%, r^{CC}=11.52%, and ρ=2.

APPENDIX TABLE 1
ROBUSTNESS: INCOME SHOCK PERSISTENCE AND BEQUESTS

	(1)	(2)	(3)	(4)	(5)
	Benchmark	$\alpha = 0.9$	$\alpha = 0.9$, Full update	Bequest x1.25	Bequest x0.75
Hyperbolic					
Parameter Estimates $\hat{\theta}$					
$\hat{\beta}$	0.7031	0.6908	0.6227	0.6964	0.6914
s.e. (i)	(0.1093)	(0.0788)	(0.1269)	(0.1168)	(0.0806)
$\hat{\delta}$	0.9580	0.9521	0.9475	0.9559	0.9581
s.e. (i)	(0.0068)	(0.0079)	(0.0122)	(0.0072)	(0.0068)
Goodness-of-fit					
$q(\hat{\theta}, \hat{\chi})$	67.2	69.5	73.2	67.3	65.9
$\xi(\hat{\theta}, \hat{\chi})$	3.01	7.40	7.64	2.54	10.2
p-value	0.222	0.025	0.022	0.281	0.006
Exponential					
Parameter Estimates $\hat{\theta}$					
$\hat{\beta}$	1.0000	1.0000	1.0000	1.0000	1.0000
s.e. (i)	-	-	-	-	-
$\hat{\delta}$	0.8459	0.8386	0.8037	0.8458	0.8459
s.e. (i)	(0.0249)	(0.0395)	(0.0528)	(0.0251)	(0.0254)
Goodness-of-fit					
$q(\hat{\theta}, \hat{\chi})$	435.6	439.6	452.1	435.5	435.8
$\xi(\hat{\theta}, \hat{\chi})$	217	340.2	369.0	211.6	216.4
p-value	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10

Source: Authors' calculations. Note: The variable α is the autocorrelation coefficient in the income process, and measures the persistence of shocks. Column 2 reports the effect of increasing α to 0.9. Column 3 differs from Column 2 in updating estimates of the other income parameters, while holding α fixed at 0.9, and also updating the variance-covariance matrix for all the first stage parameters. Columns 4 and 5 respectively increase and decrease the strength of the bequest motive by 25%.

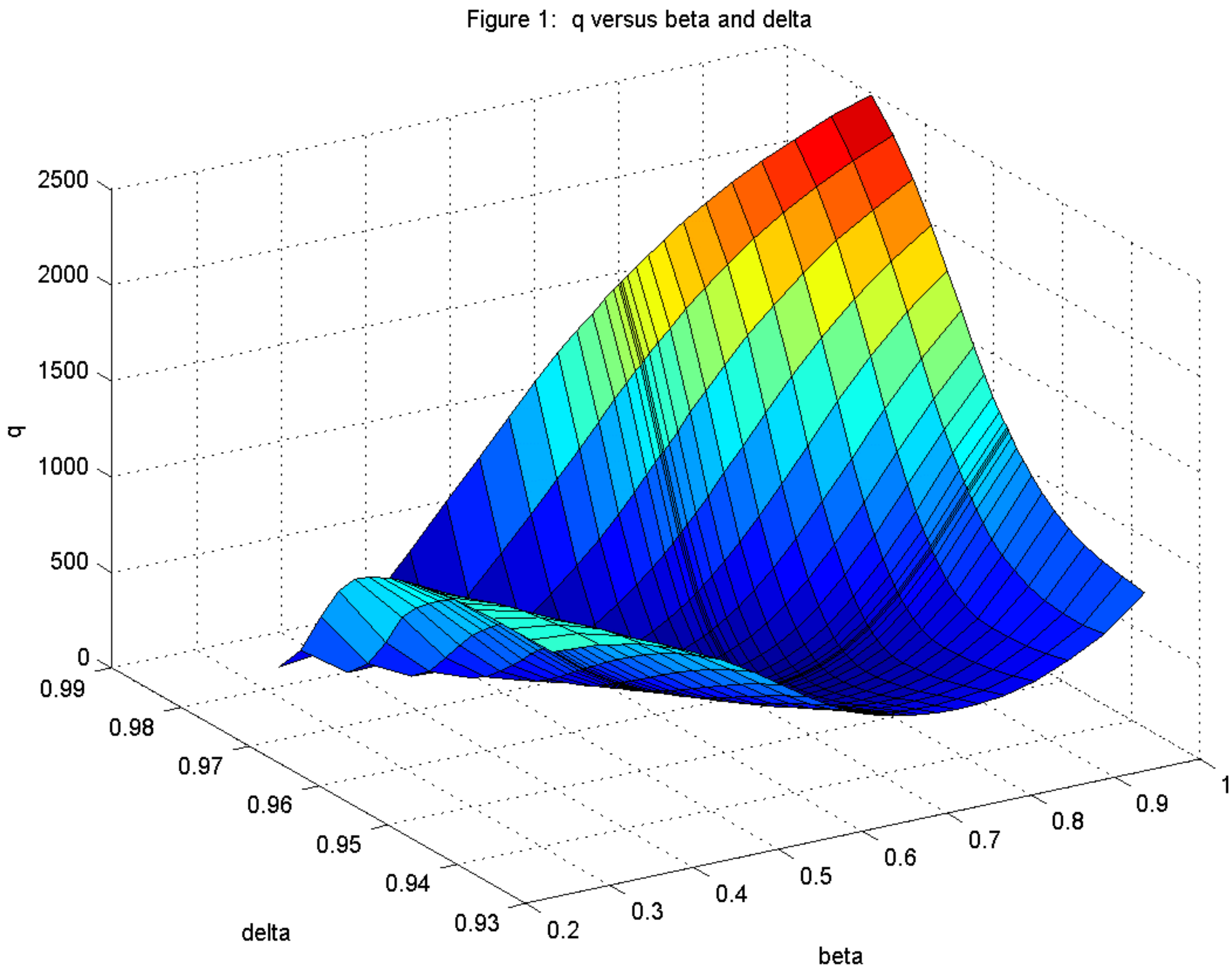

Figure 1: This figure plots the MSM objective function with respect to beta and delta under the paper's benchmark assumptions. The objective, q, equals a weighted sum of squared deviations of the empirical moments from the moments predicted by the model. Lower values of q represent a better fit of the model, and the (beta,delta) pair that minimizes q is the MSM estimator.

Notice that q appears strictly convex in the plotted region. It rises rapidly when beta and delta both rise or both fall. For a large range, q remains fairly small as beta rises and delta falls, but q rises quickly, for all values of delta, as beta approaches 1.

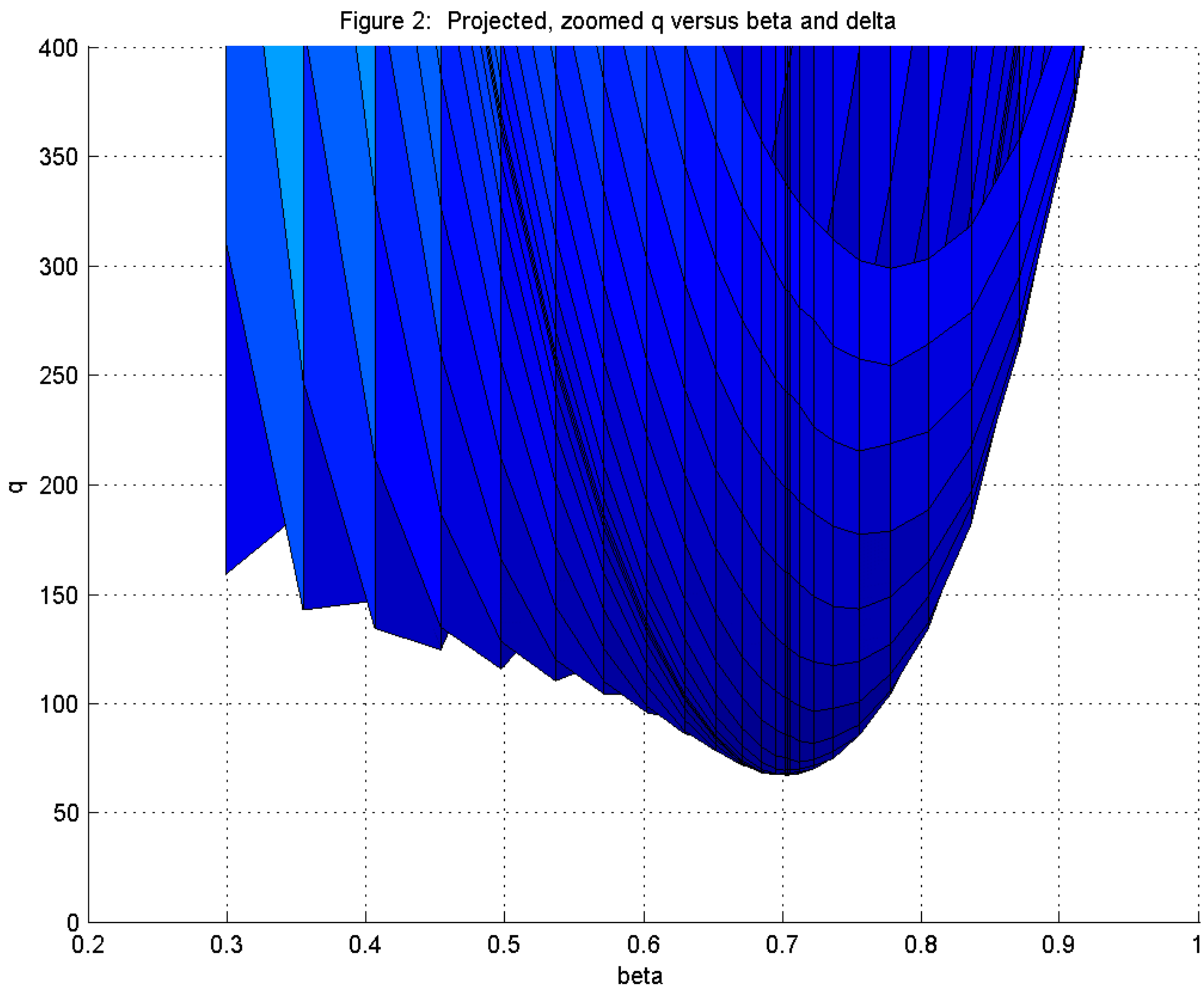

Figure 2. This figure displays the projection of Figure 1's plot of q(delta, beta) along the delta axis. It illustrates how q can remain fairly low, and hence the model's fit fairly good, for very low values of beta, while q rapidly explodes as beta approaches 1. Intuitively, at very low values of beta, higher values of delta can compensate for beta. To some extent, beta and delta are substitutes. However, as beta approaches 1 credit card borrowing declines rapidly and delta, which needs to remain fairly high to generate retirement wealth accumulation, cannot substitute for beta. The jagged appearance of the lower envelope for low beta is an artifact of Matlab's graphics routines, not an actual feature of the smooth surface.