Estimating Discount Functions from Lifecycle Consumption Choices

David Laibson, Andrea Repetto, and Jeremy Tobacman

January 26, 2004*

Abstract

This paper uses field evidence and a structural consumption-savings model to estimate discount functions. Evidence on wealth accumulation implies that people act patiently when considering long-term decisions, while data on credit card borrowing and consumption-income comovement suggests impatient behavior in the short term. Using the Method of Simulated Moments we estimate an institutionally rich model that features stochastic labor income, liquidity constraints, child and adult dependents, liquid and illiquid assets, revolving credit, retirement, and quasi-hyperbolic discount functions. We find benchmark estimates of 40% for the short-term discount rate and 4.3% for the long-term discount rate. Most specifications reject the null hypothesis of time-consistent exponential discounting. Exact quantitative results depend on assumptions about the return on illiquid assets and the coefficient of relative risk aversion.

JEL classification: D91 (Intertemporal Consumer Choice; Lifecycle Models and Saving), E21 (Consumption; Saving)

*David Laibson is Professor of Economics, Harvard University, and Faculty Research Fellow, National Bureau of Economic Research, Cambridge, Massachusetts. Andrea Repetto is Assistant Professor of Industrial Engineering, University of Chile, Santiago, Chile. Jeremy Tobacman is a graduate student in economics, Harvard University, Cambridge, Massachusetts. 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, and seminar audiences at Harvard University, Stanford University, University of Chile, Central Bank of Chile and the NBER. Laibson acknowledges financial support from the National Science Foundation (SBR-9510985), the National Institute on Aging (R01-AG-16605), the MacArthur Foundation 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.

1 Introduction

Intertemporal preferences play a key role in almost all important decisions. To measure time preference most authors have relied on laboratory studies in which the experimenter controls the choices that subjects face. Laboratory experiments typically ask subjects to weigh immediate rewards against delayed rewards. A typical study asks subjects if they would prefer $\$X$ now or $\$Y$ in the future (e.g., a month from now). Researchers impute discount rates from such tradeoffs.

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

Research using 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 databases on household behavior. However, field data are difficult to interpret since it is impossible for the researcher to know exactly what tradeoffs households actually face in real-world markets.

Given all of these considerations, laboratory and field research complement each other. Each has a useful role to play. Hence it is surprising that most of the efforts to estimate discount functions have used laboratory evidence. This imbalance is particularly true of the recent research on dynamically inconsistent time preferences. Hundreds of studies beginning with **?** and summarized in Ainslie (1992) have examined such time preferences with laboratory evidence while only a handful have attempted to do this with field data. Moreover, existing studies using field and experimental data, as surveyed by Frederick, Loewenstein and O'Donoghue (2002), have achieved little consensus about the form and amount of time discounting in economic decisions.

The current paper uses field data on lifecycle consumption choices to estimate intertemporal time preferences and formally test for dynamically inconsistent time preferences. Specifically, we use numerical methods to recursively solve and simulate a structural "buffer stock" model of lifecycle consumption and investment choices. This model includes a rich array of constraints and stochastic events that consumer face – stochastic labor income, liquidity constraints, liquid and illiquid assets, revolving credit, household dependents, and retirement –, and thus controls for a number of relevant factors that affect intertemporal decisions.

We then 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).[1] 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, which are identified from key moments in the consumption literature that characterize 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 key payoffs. First, this paper uses field data to estimate the parameters in both the dynamically consistent exponential discount function and a dynamically inconsistent alternative, the quasi-hyperbolic discount function. Second, we run formal econometric horse races between these discounting models, using both t-tests and overidentification tests. Finally, we ask whether these models accurately predict the most important behavioral regularities in the lifecycle literature.

When we adopt an exponential discount function, the MSM procedure estimates an annual exponential discount rate of 15%. By contrast, when we adopt a quasi-hyperbolic discount function, the MSM procedure estimates a short-run annualized discount rate of 40% and a long-run annualized discount rate of only 4%. All of these estimates are statistically significant. Our quasi-hyperbolic estimates imply a formal rejection of the exponential null hypothesis that the short-run discount rate is equal to the long-run discount rate.

Our overidentification tests reinforce these conclusions. Only the exponential model is con-

[1] Gourinchas and Parker (2002) and French and Jones (2001) use MSM to estimate different aspects of consumption models. Gourinchas and Parker identify the (exponential) discount rate and the coefficient of relative risk aversion off lifecycle consumption profiles. French and Jones assess 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.

sistently rejected by overidentification tests. Intuitively, the exponential model cannot simultaneously explain high levels of credit card borrowing and high levels of wealth accumulation. By contrast, the quasi-hyperbolic model implies that consumers will simultaneously act patiently and impatiently, because they have conflicting short-run and long-run discount rates.[2] In theory, low long-run discount rates explain why households accumulate substantial (illiquid) retirement wealth at interest rates of about 5%, while high short-run discount rates imply that the *same* households borrow regularly on credit cards at interest rates of 16%. By accumulating wealth in illiquid form, households commit themselves to act patiently in the future (i.e., not spending down accumulated assets). However, when liquid assets and unused credit card balances are available, households splurge whenever they can and therefore appear impatient. Consumers seem to be of two minds, acting patiently as they accumulate retirement wealth and acting impatiently in the credit card market.

We conclude the paper by reporting a wide range of robustness checks that reinforce our earlier findings, but identify the limits of our results. Most importantly we find that our estimates are relatively sensitive to assumptions about the return on illiquid assets and the coefficient of relative risk aversion.

Our quasi-hyperbolic time preference parameter estimates approximately match those of other authors who have estimated these parameters with structural models and field data. Paserman (2002) obtains identification from heterogeneity in unemployment durations and reservation wages to find estimates of a short-run discount rate that range between 10% to 60% and a long-run discount rate of 0.1%. He rejects the exponential discounting null hypothesis for two of three subsamples. Fang and Silverman (2002) estimate a model of "naive" quasi-hyperbolic discounting in which decision-makers incorrectly believe that they will have exponential discount functions in the future.[3] Using data on welfare recipients, they find a short-run discount rate of 57% and a long-run discount rate of 8%, though they cannot reject the null hypothesis of exponential discounting.

When we constrain our model to exponential discounting, we estimate a constant discount rate

[2] See Angeletos, Laibson, Repetto, Tobacman and Weinberg (2001) and Laibson, Repetto and Tobacman (2003) for expositions of this intuitive argument.

[3] See Akerlof (1991) and O'Donoghue and Rabin (1999a, 1999b) for theory on naive hyperbolic discounters. Hyperbolics who *are* aware that they will be hyperbolic in the future are called "sophisticates." 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.

of 15%. 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 require our model to fit simultaneosuly wealth accumulation data and credit card borrowing data.

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

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 consumption excess sensitivity. Table 1 summarizes these moments and Appendix 1 contains a detailed description of the data sources and estimation procedures. All of the analysis that we conduct applies to US 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).[4]

The first statistic, $\%\ Visa$, is the fraction of households who borrow on credit cards.[5] Our analysis finds that 67.8% of households pay interest on credit card debt each month. This percentage measures households who self-report that they did not pay their bill in full at the end of the last month (SCF).

We construct the second statistic, $mean\ Visa$, by dividing age-specific credit card borrowing by mean age-specific income. We then average this fraction over the lifecycle. The average household has outstanding credit card debt equal to 11.7% of the mean income of its age cohort (SCF, Fed).

The third statistic, CY, represents the excess sensitivity of consumption in response to predictable income changes. We estimate that the marginal propensity to consume is 23% of the

[4] Laibson et al. (1998, 2003) also examine households whose heads do not have high school degrees and households whose heads have college degrees, and find qualitatively similar results.

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

income change (PSID). This figure is consistent with other analyses in the literature.[6]

The final statistic, *wealth*, approximately averages the wealth-to-income ratios for households with heads aged 50-59, excluding 'involuntary' accumulation like Social Security and other defined benefit pensions. To downweight outliers we apply the arctan function to each ratio before averaging.[7] This *wealth* measure equals 2.60.

These four moments reflect important empirical characteristics of lifecycle behavior. Two additional features of the data help us interpret our results. First, most household wealth is in illiquid form. According to the SCF, adopting an expansive definition of liquidity, the average household has only 18.6% of assets in liquid form.[8] Second, though there is considerable heterogeneity among households, credit card borrowing is ubiquitous across the entire distribution of wealth. Table 2 reports the fraction of households borrowing on credit cards by age and by wealth quartile. Even among the households with a head between ages 50-59 who are between the 50th and 75th wealth percentiles, 56% did not repay their credit card in full the last time they paid their credit card bill.

3 Consumption-Savings Model

We adopt a partial equilibrium model based on the simulation literature pioneered by Carroll (1992, 1997), Deaton (1991), and Zeldes (1989), and extended by Hubbard, Skinner and Zeldes (1994, 1995), Gourinchas and Parker (2002), and Laibson et al. (1998, 2003). Our framework incorporates all of the standard features of earlier lifecycle simulation models as well as credit cards, age-dependent household size, illiquid assets, and quasi-hyperbolic time preferences. All of the model's features matter for its predictions, but we focus on the implications of Section 2's data for insight into the representation of time preferences. We divide the model summary into eight parts: demographics, non-asset income, liquid assets and noncollateralized debt, illiquid assets, budget constraints, preferences, equilibrium, and simulation. The values of our estimated inputs of the model, the first-stage parameters, can be found in Table 3.

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

[7] Our results are robust to different choices of this function.

[8] We include cash, checking and savings accounts, money market accounts, call accounts, CDs, bonds, stocks and mutual funds (as long as they are outside retirement accounts).

3.1 Demographics

We assume that independent economic life begins at age 20. Households face a time-varying, exogenous hazard rate of survival, s_t, calibrated with data from the U.S. National Center for Health Statistics (1994). To ease our computational burden, we assume that no household lives past age 90. Household composition varies deterministically and exogenously with age (calibrated from the PSID) as children and adult dependents enter and leave the household. Following Blundell, Browning and Meghir (1994) and Attanasio and Browning (1995), effective household size n_t equals the number of adults plus 0.4 times the number of children under 18. We assume households always have both a head and a spouse and calibrate the model accordingly.

3.2 Income from transfers and wages

Let Y_t represent all period t after-tax income from transfers and wages, including labor income, inheritances, private defined benefit pensions, and government transfers including Social Security. We assume labor is supplied inelastically, so Y_t is exogenous. We model $y_t = \ln(Y_t)$ during working life as the sum of a cubic polynomial in age, a Markov process η_t that approximates an underlying AR(1) process, and an iid normally distributed error. During retirement, we model y_t as the sum of a linear polynomial in age and an iid mean-zero normally-distributed error. Retirement occurs exogenously at age T. The income process and the retirement age are calibrated from the PSID.

3.3 Liquid assets and noncollateralized debt

Let X_t represent liquid asset holdings at the beginning of period t before receipt of income. If $X_t < 0$ then credit card debt was held between $t-1$ and t. We introduce a credit limit at age t of λ times average income at age, i.e., $X_t \geq -\lambda \cdot \bar{Y}_t$. The limit is calibrated from the 1995 SCF. Our model precludes consumers from simultaneously holding liquid assets and credit card debt, though such behavior has been documented among a small fraction 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 real interest rate on credit-card borrowing of R^{CC}. Our estimate of R^{CC} captures the impact of bankruptcy, default, and inflation, which lower consumers' effective interest payments, using data from the FRB, the American Bankruptcy Institute, and the CPI.

3.4 Illiquid assets

Let Z_t represent illiquid asset holdings at the beginning of period t, and assume that $Z_t \geq 0, \forall t$. Illiquid assets in our model generate two types of returns: capital gains and consumption flows. We set the gross rate of capital gains equal to $R^Z = 1$ and the annual consumption flows equal to $\gamma \cdot Z_t = 0.05 \cdot Z_t$. We assume the return on Z is considerably higher than the return on X. We adopt the assumption of complete illiquidity of Z: transaction costs are large enough that the asset can never be sold. Angeletos et al. (2001) and Laibson et al. (2003) find that assuming partial illiquidity, i.e. fixed and proportional costs of withdrawal from Z, generates similar simulations, and in the robustness section we assess the impact of making Z more attractive. This asset is obviously quite stylized, to preserve computational tractability, but we view Z as analogous to home equity. See the discussion in Appendix 2.

3.5 Dynamic and static budget constraints

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,

$$X_{t+1} = R^X \cdot (X_t + I_t^X) \tag{1}$$

$$Z_{t+1} = R^Z \cdot (Z_t + I_t^Z). \tag{2}$$

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

3.6 Preferences

We adopt constant relative risk aversion instantaneous utility functions and quasi-hyperbolic discount functions. 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. The first expression in the bracketed term is the utility flow that arises in period $t + i$ if the household survives to age $t + i$. The second expression is the termination payoffs 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 $-\ln(\delta)$. The restriction $\beta = 1$ implies time consistency and exponential discounting.

3.7 Equilibrium

Following the work of Strotz (1956) we model consumption choices as an intra-personal game. Selves $\{20, 21, ..., 90\}$ are the players in this game. Taking the strategies of other selves as given, self t picks a strategy for time t that is optimal from its perspective. This strategy is a mapping from the Markov state variables, $\{t, X+Y, Z, u\}$, to the choice variables $\left\{I^X, I^Z\right\}$. An equilibrium is a fixed point in the strategy space, such that all strategies are optimal given the strategies of the other players. We solve for the equilibrium strategies using a numerically implemented backwards induction algorithm.

Our choice of the quasi-hyperbolic discount function simplifies the induction algorithm. Let $V_{t,t+1}(\Lambda_{t+1})$ represent the time $t+1$ continuation payoff function of self t. Then its 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 $\left\{I^X, I^Z\right\}$ 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. Note that dynamic inconsistency in preferences is reflected in the fact that β appears in Equation 3 — reflecting self t's discount factor between t and $t+1$ — but does not appear in Equation 4, since self $t-1$ does not use it to discount between t and $t+1$.

Equations 3 and 4 jointly define a functional equation which is not a contraction mapping. Hence, the standard dynamic programming results do not apply to this problem. Specifically, V does not inherit concavity from u, the objective function is not single-peaked, and the policy functions are in general discontinuous and non-monotonic.[9] We have adopted a numerically efficient solution algorithm — based on local grid searches — which iterates our functional equation in the presence of these non-standard properties.[10]

[9] See Laibson (1997b).

[10] Twenty-five minutes are required on a 1.4GHz Athlon machine to solve the 70-period lifecycle problem for a single set of parameter values. Typically about 200 such solutions are required to obtain a (β, δ) parameter estimate with standard errors, so the total run-time for an estimate is about 4 days.

3.8 Simulation

We simulate the lifecycle choices of $J_s = 5000$ individual households. We generate J_s independent streams of income realizations according to the process described above. Households 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 second stage of the MSM estimation procedure described below. Note that the simulated profiles, and hence the summary 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. Imprecise replication of the exact theoretical predictions, arising from the finite size of the simulation, 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, and because MSM provides a way of accounting for additional uncertainty from simulation error.[11] The current Section describes our procedure. Appendix 3 presents derivations and some technical details.

We implement MSM in two stages. The first stage estimates nuisance parameters using standard GMM techniques. Some authors describe this as the "calibration" stage. In our case consistent estimates $\hat{\chi}$ of $N_\chi = 28$ first stage parameters χ are found, along with the variance Ω_χ of the estimate $\hat{\chi}$.[12] Our estimates in the first stage match those of other researchers. The details of this portion of the procedure were alluded to in the description of the model and can be found in Laibson et al. (2003) and in Appendix 2.

[11] 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 techniques.

[12] 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. See Table 3.

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.[13] 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. Specifically, we use Section 2's data 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. 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 $\bar{m}_{J_m}$.[14] 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 V_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 calculated directly and consistently from sample data.[15]

[13] In principle, θ and χ could be estimated simultaneously. We separate the task for three reasons (Gourinchas and Parker 2002). First, lacking good consumption/savings panel data, we turn to population aggregates to identify θ. This entails substantial loss of information. In the first stage, we are able to use detailed longitudinal household data on income and family characteristics to identify χ. Second, most of the data we use to identify θ and χ come from separate datasets and are therefore uncorrelated. Exceptions are the credit limit and CY. Covariances between the second stage moments and the credit limit we use are approximately zero, and CY's large standard error means its possible covariation with income process parameters can have little effect on the final results. Third, with current technology it is computationally infeasible to increase the number of parameters estimated using our model.

[14] Though the main text does not discuss it, the procedure accounts for the fact that J_m differs for different moments. Appendix 3 contains details.

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

Let W be a positive definite N_mxN_m weighting matrix. Define

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

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[16]

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

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 3,

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

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

This is the equation we use to calculate uncertainty in our estimates of θ. All derivatives are replaced with consistent numerical analogues, which we calculate using the model and simulation procedure.[17] We estimate Ω_g and Ω_χ consistently from sample data. After obtaining estimates using the weighting matrix $W = \Omega_g^{-1}$, we construct and use the optimal weighting matrix $W_{opt} = \left[\Omega_g + \Omega_g^s + G_\chi \Omega_\chi G_\chi'\right]^{-1}$. Many authors have found optimally-weighted GMM procedures lead to biased estimates in small samples, so we emphasize the results from using $W = \Omega_g^{-1}$.

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 arbitrarily better approximation for the true population than the sample, the simulation correction disappears. Next examine the

[16] We perform this minimization with Matlab's built-in 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 consumption functions discussed in Subsection 3.7.

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

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 in the number of observations we use to estimate $\hat{\chi}$. The first stage correction also increases with the sensitivity of the second-stage moments to changes in the first-stage parameters, G_χ.

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 arrive at the paper's three payoffs discussed in the introduction. First, we report estimates for the discount factors β and δ[18], and we also report the effect of imposing exponential discounting by requiring $\beta = 1$ (leaving δ as the only free parameter). Second, we econometrically compare the predictive power of these discounting models, using both t-tests and overidentification tests. Finally, we ask whether these models accurately predict the most important behavioral 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, as discussed in Appendix 2, these parameters

[18] First-stage results are reported in Table 3.

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. Sensitivity to the model's other parameters is already accounted for by the first stage correction. We focus on the case $W = \Omega_g^{-1}$ but 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. Globally, *wealth* increases in both β and δ while the other moments decrease in β and δ. The *wealth* variable ranges approximately from the credit limit to 10 (i.e., the maximum value permitted under the scaling transformation). The variables $\%\ Visa$ and CY range approximately from 0 to 1 over the parameter space, while $mean\ Visa$ varies from 0 to λ. In a region in the middle of the parameter space, $\%\ Visa$ and $mean\ Visa$ are essentially constant with respect to δ. Empirically, $wealth = 2.6$, $\%\ Visa = 0.68$, $mean\ Visa = 0.13$, and $CY = 0.23$, with standard errors of 0.13, 0.015, 0.01, and 0.11, respectively.

Identification of β and δ can be visualized by plotting $q(\theta, \hat{\chi})$. Recall that q is a weighted sum of squared deviations of the simulated moments from their empirical analogs. Smaller values of q reflect closer fits of the model to the data. Observe in Figure 1 that q looks like an upward-opening paraboloid. As β and δ fall, $wealth(\theta, \hat{\chi})$ nears $-\lambda$ and $\%\ Visa(\theta, \hat{\chi})$, $mean\ Visa(\theta, \hat{\chi})$, and $CY(\theta, \hat{\chi})$ all approach 1. Conversely, as β and δ approach 1, $wealth(\theta, \hat{\chi})$ nears 10 and $\%\ Visa(\theta, \hat{\chi})$, $mean\ Visa(\theta, \hat{\chi})$, and $CY(\theta, \hat{\chi})$ all fall toward zero. In these regions none of the moments can be matched and q becomes large.

Figure 1 exhibits an extended valley in the plot of q, traversing from high δ and low β to low δ and high β. Parameter pairs in this valley minimize the gap between the empirical and simulated moments. Notice that q is much more sensitive to δ than to β, because δ is exponentiated in the discount function while β is not. In addition, the orientation of the valley indicates β and δ are, to some extent, substitute forms of discounting. When δ is high, low values of β are required to match the empirical facts, and vice-versa. In particular, the valley extends to the region of very high δ and low β. However, q rises relatively quickly out of the valley as β approaches 1. As β approaches 1, the model can not match all of the empirical moments, because matching the credit card data now requires a low value of δ, but with δ so low, wealth accumulation vanishes. Figure 2 displays a higher-resolution plot, which emphasizes the point that the model cannot match the data when β is close to 1.

The figures reflect the key intuition that low long-term discount rates are necessary to match observed levels of retirement wealth, while high short-term discount rates are demanded by the data on credit card borrowing and excess sensitivity. Households will only accumulate for retirement in illiquid form at interest rates of about 5% if their long-term discount rates are less than 5%, but they will only borrow on credit cards at interest rates of about 12% if their short-term discount rates are much higher than 12%. Evidently, quasi-hyperbolic discount functions with $\beta < 1$ accomodate these seemingly dichotomous behaviors.

5.2 Benchmark Estimates

We report our findings under the benchmark assumptions in Table 4. In the quasi-hyperbolic case (Column 1) we obtain $\hat{\beta} = 0.7031$, with a standard error (s.e.(i) in the Table) of 0.1093. For this specification $\hat{\beta}$ lies significantly below 1; the t-statistic for the $\beta = 1$ hypothesis test is $t = \frac{1-0.7031}{0.1093} = 2.72$. The benchmark value for $\hat{\delta}$ corresponds to a long-term discount rate of $-\ln(0.958) = 4.3\%$, close to values estimated or adopted by other authors. We find that $\hat{\delta}$ is estimated fairly precisely, with a standard error of about seven tenths of a percent.

At the estimated parameter values, the quasi-hyperbolic model generates the moment predictions reported in Column 1 of Table 4. 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. The size of these deviations is mitigated by the considerable first stage uncertainty: high sensitivity to imprecisely-measured χ means the model is actually doing a good job. The goodness-of-fit measure $\xi\left(\hat{\theta}, \hat{\chi}\right) = 3.01$ reflects the quality of the fit, comparing favorably with 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 $\hat{\delta} = 0.8459$ – a discount rate of 16.7% – and a standard error of 0.0249. At these point estimates, the empirical facts about credit card borrowing and excess sensitivity are matched quite well. However, with such a high discount rate the model cannot account for observed wealth data. Instead, it predicts $wealth = -0.05$, since $wealth$ 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.[19] The best fit available under an exponential model predicts that typical households have negative total assets in their peak accumulation years. Goodness of fit naturally suffers: $\xi(1, 0.8459, \hat{\chi}) = 217 \gg 3.01 = \xi(0.7031, 0.958, \hat{\chi})$. Since with the exponential restriction we estimate only one parameter, we compare $\xi(1, 0.846, \hat{\chi})$ to 11.34, the 99% critical value of a chi-squared distribution with three degrees of freedom.[20] The p-value represents the probability that the benchmark model could have generated the observed data, so the overidentification test rejects the exponential model at the 1% level. More important than the exact figures, though, is the fact that the quasi-hyperbolic p-value exceeds the exponential one by many orders of magnitude.

The standard errors reported as "s.e.(i)" in Table 4 incorporate corrections for the first stage and for the simulation error. Without these corrections we find the numbers reported as s.e.(ii), s.e.(iii), and s.e.(iv) in the Table. Evidently the simulation correction matters little. Comparing s.e. (i) and (ii), we see that if the simulation were infinitely large, so that it exactly captured the properties of the theoretical population, the standard error on β would fall only from 0.1093 to 0.1090. Comparing s.e. (iii) and (iv), the simulation correction would cause an increase in the variance by about 20% if the first stage correction were irrelevant. Twenty percent is the approximate ratio of second stage moment observations J_m to simulation observations J_s, as determined by our implementation of Equation 7. However, the standard errors are strongly 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 from 0.1093 to 0.017. In other words, consistent estimates of Ω_θ

[19] 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 $wealth$ could 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 5.3.2 provide suggestive evidence.

[20] Recall that above we compared $\xi(0.705, 0.958, \hat{\chi})$ 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 an arbitrary parameter.

depend strongly on including the first stage correction. If we had not incorporated the first stage, our standard errors would have been biased down by a factor of six.

Using the optimal weighting matrix largely preserves the pattern of the benchmark results. Our optimal-weights findings are reported in Columns 4 and 5. 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 borrowing facts.

Uncertainty in all of the first stage parameters except γ, r^{CC}, and ρ has been incorporated into the standard errors reported above[21]. However, γ, r^{CC}, and ρ are difficult to pin down empirically so in the next subsection we explore the robustness of our findings to changes in those parameters.

5.3 Robustness

5.3.1 Parameter Perturbations

We begin by perturbing γ, r^{CC}, and ρ one by one from their benchmark values (i.e., $\gamma = 5\%$, $r^{CC} = 11.52\%$, and $\rho = 2$) and report the resulting estimates of β and δ in Table 5. 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 return on a marginal dollar of home equity: it equals the savings in avoided interest resulting from paying off a dollar of one's mortgage. We interpret 3.38% as being at the low end of a range of possible assumptions about returns to the illiquid asset Z. Actual returns might be higher than this net mortgage interest rate if, for example, liquidity constraints or transaction costs of loan renegotiation limit how fast people can invest in their homes.

In Column 2 we find a much lower estimate of $\hat{\beta}$ and a much higher estimate for $\hat{\delta}$ than in the benchmark case. Intuitively, when the returns on illiquidity are close to those on positive liquid

[21] Our measure of r^{CC} is constructed from aggregate data, so its true variability is underestimated in the first stage. See Appendix 2 for a discussion.

assets ($R - 1 = 2.79\%$), consumers must have significant short-run impatience (low β) to prefer the illiquid asset. Quasi-hyperbolic consumers splurge liquid savings, so they can only accumulate enough to match the *wealth* moment by saving in Z. Thus $\hat{\delta}$ rises sharply. Since $\hat{\delta}$ and $\hat{\beta}$ are substitutes globally, $\hat{\beta}$ must fall to restore the optimal fit of the other facts. In the bottom half of Column 2, we find exponential results that are identical to the benchmark case. This occurs because simulated exponential households accumulate zero illiquid wealth at the estimated $\hat{\delta}$ in both the benchmark case and in the perturbation.

In Column 3, we consider the case of $\gamma = 6.59\%$. This is the figure Flavin and Yamashita (2002) report as the total returns to housing. We view this figure as falling at the upper end of a range of possible returns to home equity. In fact, γ may be lower than 6.59% because of transaction costs of selling one's home and moving. In addition, the costs of mortgage interest payments must be netted out of the returns to find the γ of our model.[22] We now find a higher estimate of $\hat{\beta}$ and a lower estimate for $\hat{\delta}$ than in the benchmark. As γ rises to approach the credit card interest rate, the model can accomodate simultaneous illiquid wealth accumulation and credit card borrowing more easily, even for more time-consistent consumers. Despite the increased estimate $\hat{\beta}$, the smaller standard error means the $\beta = 1$ hypothesis is rejected at the 99% confidence level.

The comparative statics of increasing γ contrast with the less powerful perturbation of decreasing r^{CC}.[23] In Column 4 we assume $r^{CC} = 10\%$ and find that $\hat{\beta}$ rises and $\hat{\delta}$ falls relative to the benchmark case. The standard errors change little. 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 equals the debt-weighted average interest rate from the FRB, minus inflation, minus the personal bankruptcy rate. This ignores (i) the fact that the marginal utility of consumption may be especially high in the bankruptcy state, implying an underestimate of the correction, and (ii) that bankruptcy carries stigma, implying an

[22] If on average half the mortgage has been paid off, then perhaps the appropriate returns to Z are 6.59%-$\frac{1}{2}$3.38% = 4.9%, which is close to our benchmark assumption of $\gamma = 5\%$.

[23] Perturbations to γ influence parameter estimates more than perturbations to r^{CC} because asset accumulation is higher than borrowing.

overestimate of the correction. We favor the middle specification as our benchmark, but observe that changes in r^{CC} of 150 basis points reported in Columns 4 and 5 have little quantitative effect on the time preference estimates.

Finally, we examine the effect of varying the coefficient of relative risk aversion ρ. Economists disagree notoriously about how to calibrate ρ. In order to account for the equity premium puzzle, the consumption CAPM requires $\rho > 25$ (Kocherlakota 1996). Most consumption simulation papers assume $\rho \in [.5, 5]$, consistent with typical introspective choices about hypothetical gambles in the positive domain. When liquidity constraints do not bind, ρ is the inverse of the elasticity of substitution. However, Euler Equation estimates of the EIS range roughly between 0 and 1. Using a structural approach, Gourinchas and Parker (2002) identify ρ from lifecycle consumption profiles. For different specifications they find ρ between 0.2 and 5, with a precise benchmark estimate of 0.51.

In addition to this empirical ambiguity, 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 (2002) proposes that consumption commitments could cause different local and global levels of risk aversion. Rather than entering this debate, above we adopted $\rho = 2$ for our benchmark. We now examine the effect of perturbations to $\rho = 1$ and $\rho = 3$. Column 6 of Table 5 reports the effect of assuming $\rho = 1$. We find that $\hat{\beta}$ and $\hat{\delta}$ both rise relative to the benchmark. In this specification, $\hat{\beta}$ is marginally significantly different from 1. When $\rho = 3$, as reported in Column 7, we find that $\hat{\beta}$ and $\hat{\delta}$ both fall relative to the benchmark. Goodness-of-fit worsens (i.e., $q\left(\hat{\theta}, \hat{\chi}\right)$ rises) as ρ rises. Intuitively, as ρ rises illiquid wealth accumulation and credit card borrowing both become less attractive. Illiquid assets cannot be used as a buffer when bad shocks raise the marginal utility of consumption, and anticipation of high future marginal utility reduces the temptation to splurge today at high interest rates. Risk aversion is discussed further in the Extensions subsection below.

5.3.2 Extreme Cases

We also consider two extreme cases. In Case A, we consider the effect of simultaneously assuming $\gamma = 3.38\%$, $r^{CC} = 13\%$, and $\rho = 3$, and in Case B we assume $\gamma = 6.59\%$, $r^{CC} = 10\%$, and $\rho = 1$. We consider both sets of assumptions implausible, but we present estimates for them to indicate rough upper and lower bounds on estimates of β and δ.

For Cases A and B, $q(\theta, \hat{\chi})$ exhibits features similar to the benchmark. Though the upward-opening paraboloid shifts with the assumptions about γ, r^{CC}, and ρ, q rises quickly for high-β, high-δ pairs and low-β, low-δ pairs; q has a valley proceeding from the high-β, low-δ region to the low-β, high-δ region; and the minimum of the function lies away from $\beta = 1$.

Case A combines three perturbations of first stage parameters that we saw above, in Table 5. Each of those perturbations lowered estimates of β, and in Column 1 of Table 6 we see that their combined effect results in $\hat{\beta} = 0.375$. Again the exponential model, where β is restricted to equal 1, results in a very low estimate of $\hat{\delta}$. With such a low $\hat{\delta}$, we see in Column 2 that the model predicts credit card borrowing and excess sensitivity quite well, but predicts approximately no wealth accumulation.

In Case B we combine the three other perturbations from Table 5 to find $\hat{\beta} = 0.9075$. This estimate carries a small standard error, implying that even under aggressive assumptions about γ, r^{CC}, and ρ, $\hat{\beta}$ is significantly less than 1. In the exponential case in Column 4 we find a much higher estimate of $\hat{\delta}$. Here the credit card and consumption-income comovement facts are still matched reasonably well, and the model predicts nearly realistic levels of retirement wealth. Goodness-of-fit improves (ξ falls) for both the hyperbolic and exponential specifications in Case B, relative to Case A, as well.

6 Extensions

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

6.1 The coefficient of relative risk aversion

Given the sensitivity of the model's quantitative results to the value of ρ, it would be interesting to estimate ρ simultaneously with β and δ. We anticipate two characteristics of such an exercise.

First, if $\rho = 0$ is inserted into the exponential Euler Equation, we would impute a discount rate equal to the credit card interest rate. This discount rate would be higher than what other authors assume, and surely too high to generate observed levels of wealth accumulation. Higher short-term than long-term discount rates would likely still be needed to account for the facts.

Second, the curse of dimensionality implies that estimating ρ in addition to β and δ would be computationally costly. Though we expect β would be identified away from 1, β and ρ may otherwise only be weakly separately identified. Thus we might need a very large number of function evaluations to find the minimum of $q(\beta, \delta, \rho; \hat{\chi})$.

6.2 Naivete

Our equilibrium definition adopts the standard economic assumption of unlimited problem-solving sophistication. The consumers in our model solve perfectly a complex backwards induction problem when making their consumption and asset allocation choices. Though we view this assumption as a reasonable starting point, we are not fully satisfied with it. One alternative is the model of "naif" behavior first proposed by Strotz (1956) and more recently studied by Akerlof (1991) and O'Donoghue and Rabin (1999a, 1999b). These authors propose that decision makers with dynamically inconsistent preferences make current choices under the false belief that later selves will act in the interests of the current self. Angeletos et al. (2001) find that naive and sophisticated quasi-hyperbolics similarly in many respects in lifecycle consumption models.

However, two puzzles remain which perhaps a model of partial naivete could address. First, the sophisticated quasi-hyperbolics in these simulations would be better off if they had no access to credit cards throughout their lifecycles. 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 leads naturally to the question of why quasi-hyperbolic consumers do not in fact cut up their credit cards. Second, the spread between the cost of funds and the credit card interest rate is "too high." It cannot be accounted for by standard explanations like default probabilities (Ausubel 1991). More generally, the interest rate structure, including the ubiquitous teaser rates, clamors for an explanation. A model of naive overoptimism about future impatience might resolve these questions.[24]

6.3 Heterogeneity

A third natural direction for future work would be to relax the representative agent assumption and consider population heterogeneity. 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 it does not explain why the median household both borrows on its credit cards and invests in illiquid assets. Table 2 indicates that credit card borrowing is pervasive across the entire wealth distribution.

6.4 Institutional assumptions

It would be natural to allow households to declare bankruptcy, relax the assumption that the Z asset is perfectly illiquid, and consider the option to purchase Z with collateralized debt. However, these computationally-costly modifications have been shown by Laibson et al. (2003) to have little effect on the predictions of similar models. The current paper incorporates bankruptcy through the credit card interest rate as a reduced-form. Further discussion of the characteristics of Z can be found in Appendix 2.

6.5 Second stage moment sets

Estimates of discount functions could perhaps be refined by identifying off different sets of empirical facts. One could try to identify the parameters with higher-resolution information derived from fewer observations, like wealth-to-income ratios at every age. That would provide a stricter test for both the exponential and quasi-hyperbolic models. We could also analyze the consumption drop at retirement documented by Banks, Blundell and Tanner (1998) and Bernheim, Skinner and Weinberg (1997). Both the exponential and quasi-hyperbolic benchmark models, evaluated at the estimated parameters, predict a consumption drop of 20% in a 4-year window around retirement.

In addition, we find in the SCF that households hold almost all of their wealth in illiquid form. Even with an expansive definition of liquid assets, only 18.6% of total US household wealth

[24] See also the very interesting experimental results of Ausubel (1999) and the theoretical contribution of DellaVigna and Malmendier (2001).

is liquid.[25] The exponential model overpredicts this share, while the quasi-hyperbolic model overpredicts it by a smaller – but still very large – amount.[26] Intuitively, quasi-hyperbolics tend to splurge liquid assets more often than exponential households.

Dynan (1993) finds that estimates of the coefficient of relative risk aversion using second-order log linearizations of the Euler Equation are negative and hence anomalously low. Laibson et al. (1998) show that these results could be explained by omitted variables bias: an omitted term derived from an expansion of the quasi-hyperbolic Euler Equation correlates negatively with the second-order term in the exponential Euler Equation linearization. Values of ρ imputed through these regressions could be used to identify β.

Also, the calibrated quasi-hyperbolic model predicts a small consumption boom early in life when credit cards are acquired. One could search for this in the data and try to compare with the model. Finally, many facts about the behavior of the elderly remain unexplained. Economic models that do not impose perfect illiquidity of assets have difficulty explaining why the elderly do not decumulate, borrow extensively, and/or default on their debt.

7 Conclusion

This paper uses field evidence to contribute to the discussion about the form and amount of time discounting. Central facts from the consumption literature identify the quasi-hyperbolic discounting parameters β and δ in an institutionally rich lifecycle model. U.S. households accumulate large stocks of wealth before retirement, borrow frequently and extensively on credit cards, and exhibit excess sensitivity of consumption. These phenomena can be explained best in our benchmark specification when $\beta = 0.703$ and $\delta = 0.958$. Intuitively, low long-term discount rates account for observed levels of (illiquid) wealth accumulation, and high short-term discount rates explain observed levels of credit card borrowing and excess sensitivity. The MSM procedure rejects the exponential null hypothesis and usually fails to reject the overidentification restrictions on the quasi-hyperbolic model.

[25] The methodology for calculating the share is analogous to that for calculating % *Visa*, *mean Visa*, and *wealth* described in Appendix 1. The definition includes cash, checking and savings accounts, money market accounts, call accounts, CDs, bonds, stocks and mutual funds.

[26] The calibrated hyperbolic model predicts a share equal to 37%.

Exact quantitative results are still somewhat sensitive to the specification, and some overidentification tests fail. However, the evidence reported here suggests that consumption-savings models can perform better vis à vis the data when they incorporate quasi-hyperbolic discount functions, and discount rates in the short run exceed discount rates at longer horizons. Estimation methods like those employed in this paper can help evaluate the implications of field evidence for deep behavioral parameters.

Appendix 1

We now discuss the procedures used to construct the second-stage moments. We used the SCF to derive *wealth*, % $Visa$, and *mean* $Visa$, and the Panel Study of Income Dynamics (PSID) to construct CY. Procedures are very similar for the share of liquid assets, an SCF moment described in the discussion section.

SCF Moments

We use the 1983, 1989, 1992, 1995, and 1998 SCF's to compute the *wealth* moment. We derive % $Visa$ and *mean* $Visa$ from the 1995 and 1998 SCF's. 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. The procedure is this:

For each variable of interest x first use weighted least squares, using the SCF population weights, 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.

Next, we define the "typical" household to be identical to the simulated household (i.e. with two heads, exogenous age-varying numbers of children and adult dependents, an average cohort effect, and an average unemployment effect[27]). Then for each variable we create a new variable $\widehat{x}_i$

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

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 aged between 50 and 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 calibrated 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 max(0, value of home − mortgages − 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[28].

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 confounded by the fact that aggregate credit card borrowing data from the Federal Reserve Board indicate that 1995 and 1998 SCF borrowing magnitudes are

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

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

biased downward 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 US 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, 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 $debt_i/\overline{y}$ [29]. 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 on 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.

[29] 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.

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. 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 overidentification test does not reject this specification. The vector X_{it} includes age, cohort, and business cycle effects, the change in effective family size, the mortality rate, and lagged wealth. 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.

Appendix 2

Most details of the first stage estimation are standard and exactly follow Laibson et al. (2003). Additional notes are included here.

1. Our assumption about the credit card interest rate ignores several offsetting effects. One might believe the effective rate would be lower for two reasons. Without declaring bankruptcy, households might be able to default. In addition, consumption may be unusually low in the bankruptcy state, causing the marginal value of un-repaid dollars to be unusually high. Conversely, the model does not account for the substantial stigma associated with bankruptcy (Gross and Souleles 2002b) or for the cost of future exclusion from credit markets. We feel that the estimated

interest rate does reflect the typical cost of credit card borrowing. To check for robustness, we also perform the estimation under the assumptions $r^{cc} = 10\%$ and $r^{cc} = 13\%$.

2. Three main issues arise when interpreting the illiquid asset Z. First, the specification is quite stylized, but Z shares some similarities with home equity. 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 hence 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 corrected for the tax deductibility of interest payments and inflation. If we assume $\eta \approx \gamma$, the net benefit to the homeowner is $\gamma H - \eta M. \approx \gamma(H - M) = \gamma Z$.

Second, the assumption of total illiquidity increases the motive for credit card borrowing in both the exponential and quasi-hyperbolic models. When illiquid assets have favorable returns but are highly illiquid, consumers will want to hold those illiquid assets and use credit card borrowing to smooth consumption volatility due to high frequency shocks in the income process.

Finally, for quasi-hyperbolic consumers even small delays between requests for liquidity (i.e., applications for home equity lines) and actual access to liquidity (i.e., approval of applications and release of funds) have a behavioral effect equivalent to total illiquidity (Laibson 1997a). In other words, even fairly small actual or perceived transaction costs involved in extracting liquidity from home equity will deter quasi-hyperbolics from applying for home equity loans. Hence, the extreme illiquidity assumption is particularly appropriate for the quasi-hyperbolic model.

3. To obtain an estimate for R, we take a 'typical' value from a standard series. We do not necessarily think AAA municipal bonds are the relevant saving margin for most consumers. However, other authors in the simulation literature assume values between $R = 1$ and $R = 1.05$, giving us a reasonable benchmark. The choice of R is plausibly bounded below by the aftertax real risk free rate and above by the aftertax real return to an equity index. We assume returns are certain, so if you believe in perfect markets and no behavioral explanations for the equity premium, we should adopt the risk-free rate of return. Since it is not obvious what alternative is best, we adopt a procedure that follows precedent and yields an intermediate value. In addition, for $R - 1 \;.\; \gamma - .01$, the parameter estimates are not very sensitive to the choice (though goodness of fit is somewhat sensitive). The more important margin is the difference between γ and R^{CC}.

Appendix 3

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 expectations. 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

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

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,

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

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: $\sqrt{\hat{J}_m} g(\theta_0;\chi_0) \rightarrow$

$N(0, V_g)$. We estimate $\Omega_g = \frac{V_g}{J_m} = E\left[g(\theta_0; \chi_0)\, g(\theta_0; \chi_0)'\right]$ directly from its sample counterpart.[30] 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 second term is also asymptotically normal (as the size of the simulated sample goes to infinity) with mean 0 and variance $\hat{J}_m \frac{V_g}{J_s}$. Finally, since variation in the simulation and the data are independent, $\sqrt{\hat{J}_m} g_{J_s}(\theta_0; \chi_0) \to N\left(0, \left(1 + \frac{\hat{J}_m}{J_s}\right) V_g\right)$. 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 V_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 $\sqrt{\hat{J}_m}\left(\widehat{\theta} - \theta_0\right) \to N(0, V_\theta)$, where Equation 9 implies

$$V_\theta = \left(G'_\theta W G_\theta\right)^{-1} G'_\theta W \left[\left(1 + \frac{\hat{J}_m}{J_s}\right) V_g + \hat{J}_m \cdot G_\chi \Omega_\chi G'_\chi\right] W G_\theta \left(G'_\theta W G_\theta\right)^{-1}, \qquad (10)$$

by the asymptotical 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.

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

References

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

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

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

Attanasio, Orazio and Martin Browning, "Consumption over the Life Cycle and over the Business Cycle," *American Economic Review*, December 1995, *85* (5), 1118–1137.

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

——, "Adverse Selection in the Credit Card Market," 1999. University of Maryland mimeo.

Banks, James, Richard Blundell, and Sarah Tanner, "Is There a Retirement Savings Puzzle?," *American Economic Review*, September 1998, *88* (4), 769–788.

Bernheim, B. Douglas, Jonathan Skinner, and Steven Weinberg, "What Accounts for the Variation in Retirement Wealth Among US Households?," 1997. Stanford University mimeo.

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

Blundell, Richard, Martin Browning, and Costas Meghir, "Consumer Demand and the Life-Cycle Allocation of Household Expenditures," *Review of Economic Studies*, 1994, *61*, 57–80.

Browning, Martin and Annamaria Lusardi, "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, December 1996, *34*, 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.

Chetty, Raj, "Consumption Commitments, Unemployment Durations, and Local Risk Aversion," 2002. Harvard University mimeo.

Deaton, Angus, "Saving and Liquidity Constraints," *Econometrica*, 1991, *59*, 1221–1248.

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

DellaVigna, Stefano and Ulrike Malmendier, "Contract Design and Self-Control: Theory and Evidence," December 2001. Harvard University mimeo.

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

Dynan, Karen, "How Prudent are Consumers?," *Journal of Political Economy*, 1993, *101* (6), 1104–1113.

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," April 2002.

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 and John Jones, "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," in "Mimeo" 2001.

Gourinchas, Pierre-Oliver and Jonathan Parker, "Consumption over the Life Cycle," *Econometrica*, January 2002, *70* (1), 47–90.

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), 149–186.

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

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*, 59–125.

___, ___, 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*, 42–71.

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

___, "Hyperbolic Discount Functions and Time Preference Heterogeneity," in "Mimeo" Harvard University 1997.

___, Andrea Repetto, and Jeremy Tobacman, "Self-Control and Saving for Retirement," *Brookings Papers on Economic Activity*, 1998, (1), pp. 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.

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 R 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.

Paserman, Daniele, "Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation," 2002. Hebrew University mimeo.

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

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

Strotz, Robert, "Myopia and Inconsistency in Dynamic Utility Maximization," Review of Economic Studies, 1956, 23, 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.

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. The variables are defined as follows: %Visa is the fraction of US households borrowing and paying interest on credit cards (1995 and 1998 SCF); meanVisa is the average amount of credit card debt as a fraction of the mean income for the age group (1995 and 1998 SCF, weighted by Fed aggregates); CY is the marginal propensity to consume out of anticipated changes in income (1978-92 PSID); and wealth is the weighted average wealth-to-income ratio for households with heads aged 50-59 (1983-1998 SCF).

TABLE 2

FRACTION OF HOUSEHOLDS BORROWING ON CREDIT CARDS ACROSS THE DISTRIBUTION OF WEALTH

	Wealth Distribution Percentile			
Age Group	Less than 25	25-50	50-75	Over 75
20-29	0.89	0.78	0.82	0.75
30-39	0.92	0.83	0.82	0.63
40-49	0.85	0.79	0.74	0.49
50-59	0.80	0.73	0.56	0.41
60-69	0.64	0.40	0.27	0.23
70+	0.47	0.29	0.11	0.12

Source: Authors' calculations based on the 1995 and 1998 SCF's.

TABLE 3

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-09	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

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

$\xi_t = \mu_t + \upsilon_t = \alpha\mu_{t-1} + \varepsilon_t + \upsilon_t$

β_0	β_1	β_2	β_3	β_4	β_5	β_6	α	σ^2_ε	$\sigma^2_{\upsilon\phi}$
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

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

β_0	β_1	β_2	β_3	β_4	σ^2_ξ
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: Author's estimation based on data from the PSID, SCF, FRB, and American Bankruptcy Institute.

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 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; see, however, the section on Robustness.

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

TABLE 4
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 5
ROBUSTNESS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Benchmark	$\gamma = 3.38\%$	$\gamma = 6.59\%$	$r^{CC} = 10\%$	$r^{CC} = 13\%$	$\rho = 1$	$\rho = 3$
Hyperbolic							
Parameter Estimates $\hat{\theta}$							
$\hat{\beta}$	0.7031	0.5071	0.8024	0.7235	0.6732	0.8186	0.5776
s.e. (i)	(0.1093)	(0.0441)	(0.0614)	(0.1053)	(0.1167)	(0.0959)	(0.1339)
$\hat{\delta}$	0.9580	0.9731	0.9425	0.9567	0.9595	0.9610	0.9545
s.e. (i)	(0.0068)	(0.0188)	(0.0093)	(0.0071)	(0.0045)	(0.0037)	(0.0096)
Goodness-of-fit							
$q(\hat{\theta},\hat{\chi})$	67.2	108.4	49.7	64.1	70.7	63.0	67.7
$\xi(\hat{\theta},\hat{\chi})$	3.01	16.79	5.27	12.09	10.97	7.97	1.85
p-value	0.222	0.0002	0.0717	0.0024	0.0041	0.0186	0.3965
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	0.8924	0.7841
s.e. (i)	(0.0249)	(0.0249)	(0.0250)	(0.0267)	(0.0262)	(0.0204)	(0.0357)
Goodness-of-fit							
$q(\hat{\theta},\hat{\chi})$	435.6	435.6	435.6	434.7	436.6	438.1	435.5
$\xi(\hat{\theta},\hat{\chi})$	217	217	263	177	339	349	310
p-value	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10	<1e-10

Source: Authors' calcuations.
Note: The benchmark assumes gamma=5%, rcc=11.52%, and rho=2. Columns (3) through (8) perturb parameters one at a time.

TABLE 6

EXTREME CASES

	Case A		Case B	
	(1) Hyperbolic	(2) Exponential	(3) Hyperbolic	(4) Exponential
Parameter Estimates $\hat{\theta}$				
$\hat{\beta}$	0.3750	1.0000	0.9075	1.0000
s.e. (i)	(0.4859)	-	(0.0285)	-
$\hat{\delta}$	0.9717	0.7695	0.9434	0.9359
s.e. (i)	(0.0228)	(0.0262)	(0.0059)	(0.0071)
Second-stage moments				
% Visa	0.650	0.680	0.643	0.506
mean Visa	0.188	0.153	0.155	0.097
CY	0.504	0.297	0.230	0.141
wealth	2.55	-0.06	2.62	2.52
Goodness-of-fit				
$q(\hat{\theta}, \hat{\chi})$	106.1	436.1	38.9	145.2
$\xi(\hat{\theta}, \hat{\chi})$	16.06	319.5	7.52	19.68
p-value	0.0003	<1e-10	0.0233	0.0002

Source: Authors' calculations.
Note: Case A assumes gamma=3.38%, rcc=13%, and rho=3. Case B assumes gamma=6.59%, rcc=10%, and rho=1. The benchmark assumes gamma=5%, rcc=11.52%, and rho=2.

Figure 1: q versus beta and delta

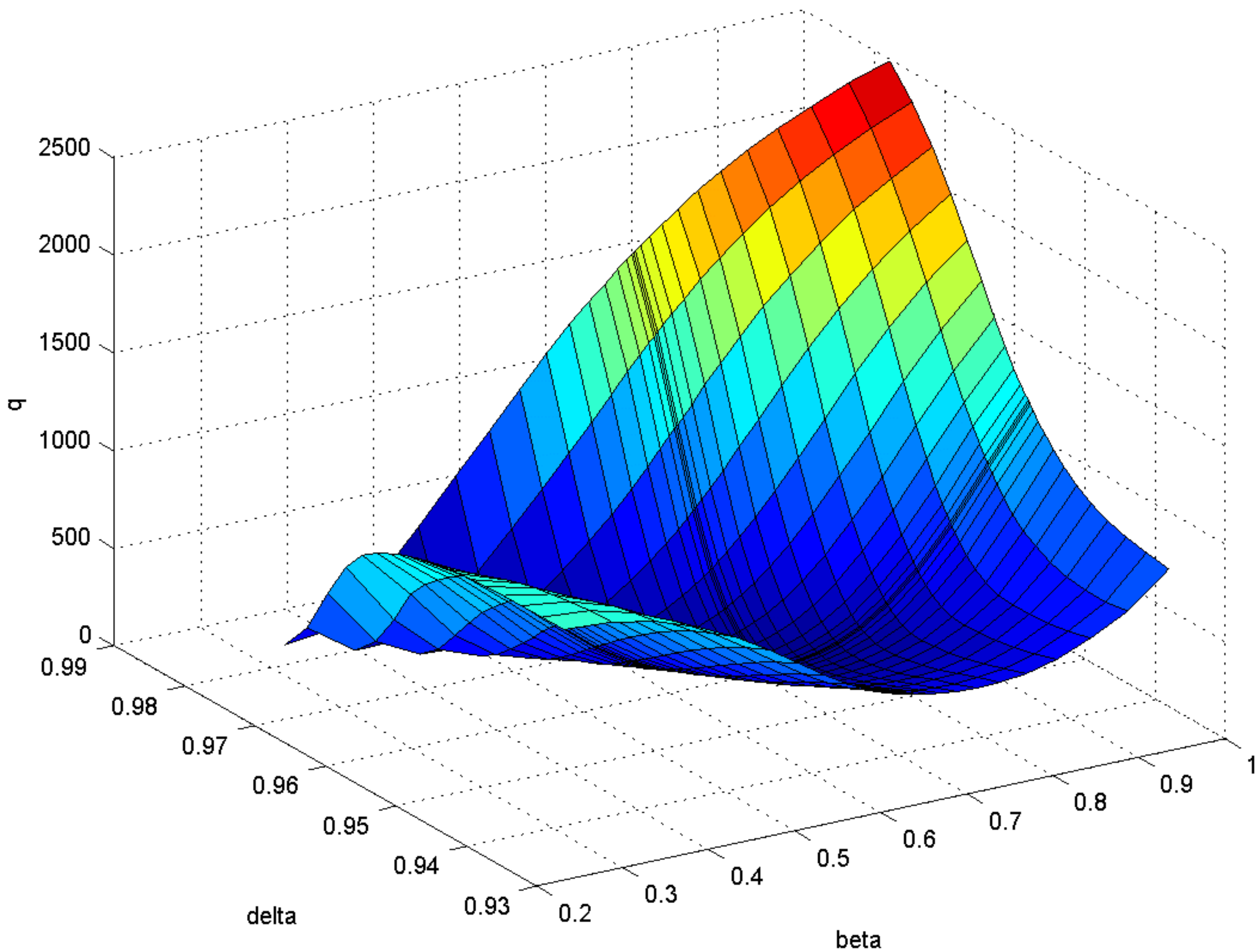

Figure 2: Projected, zoomed q versus beta and delta