Internet Appendix to "Exponential Growth Bias and Household Finance"*

Appendix A. Evidence on Exponential Growth Bias

Exponential growth bias has its intellectual origin in papers by Wagenaar and Sagaria (1975) and Wagenaar and Timmers (1978, 1979). One motivation for the studies is a Chinese parable of the "pond and duckweed" describing how even a very experienced mandarin underestimates how quickly his pond will be covered by duckweed that doubles in size every five years.

Exponential growth bias is typically measured by asking subjects in laboratory experiments or surveys to extrapolate an exponential series of the general form:

$$y = f(x) = a^{bx}$$

Typically the problem focuses on extrapolation over time, where x above becomes t, the number of periods. The context of the problem varies; examples include forecasting population, pollution (Wagenaar and Sagaria (1975)), duckweed (Ebersbach and Wilkening (2007), Wagenaar and Timmers (1979)), prices (Jones (1984), Kemp (1984), Keren (1983)), and others. Studies have varied the mode of data presentation (numerical, mathematical, or visual) and the format of questions; e.g., a study might show respondents how quickly the number of marbles grows in a jar, then ask how long it would take for the number of marbles to double or reach some other figure. Other studies graph an exponential function, then ask respondents to extend it by sketching the next few points. Most of the research focuses on *intuitive extrapolation* that does not rely on calculators; later work investigates how decision aids such as calculators or heuristics improve responses (Arnott (2006), Arnott and O'Donnell (1997)).

The central finding of this research is that individuals persistently and substantially underestimate exponential growth: they display exponential growth bias. The result is general and robust to different contexts and presentations. The magnitude of underestimation appears to be essentially orthogonal to the context of the problem, the way the data are presented (numerically, mathematically, or visually), or the frame/format of the question.

The cognitive source of exponential growth bias seems to be a strong tendency for the brain to linearize functions when extrapolating or forecasting. This tendency causes particularly large errors when the data-generating process is exponential. It causes both persistent underestimates of growth and persistent overestimates of declining series. Kemp (1984), for example, finds that consumers' recollections of (actual) past prices are persistently too high. Whether this tendency to linearize is innate or learned is an open question.

---

*Citation format: Stango, Victor, and Jonathan Zinman, 2009, Internet Appendix for "Exponential Growth Bias and Household Finance," Journal of Finance 64, pages 2807-2849, http://www.afajof.org/IA/2009.asp#dec09. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

Ebersbach and Wilkening (2007) find that younger children display greater exponential growth bias but note that schooling orients children toward linear approximation. They discuss other work finding that exponential growth bias increases in early years of schooling, and then falls.

Whether the process that respondents actually use to forecast exponential growth can be identified is also an open question. Much of the early work uses responses to fit general equations of the following form:

$$y = f(x) = \alpha \cdot a^{\beta bx}$$

Given a true data-generating process with $[\alpha, \beta] = [1,1]$ it is then possible to explicitly test for correct extrapolation of exponential growth. All studies we know that perform this test strongly reject correct extrapolation. Estimates of the coefficients vary, but in many cases respondents underestimate the exponent by a factor of ten $(\beta = 0.10)$. It also appears that many respondents linearly compensate for underestimation of the exponent by inflating the scale term $(\alpha > 1)$. Jones (1984) argues that a polynomial specification fits the data better than the one above, while Keren (1983) correctly notes that the true function used by respondents for extrapolation is unidentified, and that the goal should be parsimonious description of the data rather than identification of the true data-generating process.

Despite the robust finding that exponential growth bias exists and is systematic, there has been relatively little work exploring its economic implications. This is a bit puzzling given how direct the application is, particularly for intertemporal choice. Exceptions include Keren (1983) and Jones (1984), who both find systematic underestimation of future price increases; Kemp (1984), who finds that consumers systematically overstate past price levels; and Eisenstein and Hoch (2005), which we discuss in the text.

Appendix B. The Mathematics of Exponential Growth Bias, Future Value Bias, and Payment/Interest Bias

In this section we examine how exponential growth bias affects perceived returns to saving and perceived loan interest rates. The central question is whether a general formulation of EG bias has unambiguous predictions about how consumers perceive interest rates and returns to saving, when making inference using information commonly available in the market. We also ask how such inference changes as the time horizon of saving/borrowing changes. The analysis is relatively straightforward for returns to savings (and would not require a mathematical appendix) but is more complicated for borrowing costs. One complication is that there is no closed-form solution for the interest rate on an installment loan. This necessitates the use of tools for comparative statics on solutions to problems that are defined implicitly. Second, because we are interested in comparative statics over large ranges of the data, and because the function defining the interest rate may be highly nonlinear and/or may have multiple solutions, we use the tools of monotone comparative statics rather than the implicit function theorem to obtain our results.[1]

We begin by laying out the mathematics of borrowing and savings calculations in Section A. We then formally define EG bias in Section B. Section C presents the (relatively straightforward) proof that EG bias implies underestimation of future values on savings and that underestimation is more severe at long time horizons. Section D proves that EG bias implies payment/interest bias, and shows conditions under which payment/interest bias is more severe on short-term debt.

A. Financial Calculations and Exponentiation

Both the formula for an installment loan interest rate and the formula for a future value contain exponential growth terms. Consider first a consumer attempting to calculate the return to saving, in dollar terms, based on a current dollar value and a given interest rate. The formula for the future value FV of an amount PV saved at an interest rate i for t periods is the following exponential function:

$$FV = PV(1+i)^t$$

A similar term enters the calculation a consumer must make when inferring an interest rate from an installment loan payment m to an interest rate, maturity t and loan principal L:

$$m = Li + \frac{Li}{\left[(1+i)^t - 1\right]}$$

[1] The introduction in Milgrom and Roberts (1994) discusses why monotone comparative statics are often more desirable than the comparative statics obtained via the implicit function theorem.

Note that both expressions contain an exponential growth term $f(i,t)=(1+i)^t$, which is a specific parameterization of the general exponential function $y=f(x)=a^{bx}$ examined in the exponential growth bias literature (see Appendix A).

In what follows, we ask how underestimation of the exponential growth term $f(i,t)=(1+i)^t$ only affects inference about future values and loan interest rates. While it is possible that consumers' inference involving operations other than exponentiation could be biased as well, we focus on EG bias for two reasons. First, there is substantial evidence from cognitive psychology that errors on exponential calculations are biased, but little evidence of bias on simpler operations such as multiplication or division. Second, focusing on EG bias is useful because it provides tests that link payment/interest bias to behavior regarding savings calculations (in which the only complex mathematical operation is an exponential term).

B. Defining Exponential Growth Bias

We define exponential growth bias as a parameter $\theta \in [0,1]$ that produces underestimation of the exponential function above. Define $f(i,t,\theta)$ as a function describing a consumer's potentially biased perception of exponential growth, with the following properties:

(a) $f(i,t,0)=(1+i)^t$

(b) $f(i,t,1)>1$

(c) $f(i,t,\theta)$ is strictly decreasing in θ

(d) $f(i,t,\theta)$ is strictly increasing in i

(e) $f(i,t,\theta)$ is strictly increasing in t

(f) (increasing differences in t) for any $t'>t$ and values $[i',i'',\theta',\theta'']$ such that $f(i'',t,\theta'')-f(i',t,\theta')>0$, it is also true that $f(i'',t',\theta'')-f(i',t',\theta')>f(i'',t,\theta'')-f(i',t,\theta')$.

Property (a) defines an unbiased consumer with $(\theta=0)$ as one who correctly calculates the exponential growth term. Property (b) places a bound on the inference displayed by a biased consumer; even a consumer with maximal bias will never perceive a future value of savings as less than the present value.[2] Property (c) states that increases in bias on the unit interval reduce perceived exponential growth. Properties (d) and (e) maintain the standard assumptions that for

[2] We restrict consideration to situations where the interest rate is positive and the number of time periods is greater than one.

any degree of bias, perceived exponential growth is increasing in the interest rate and time. Property (f) states that differences in perceptions for pairs $[i', i'', \theta', \theta'']$ do not become less severe over longer time horizons. This is satisfied if $f(i,t,\theta)$ is an exponential function in *t,* meaning that our proof applies when consumers underestimate exponential growth using a more slowly growing exponential function. It is also satisfied when consumers use simpler non-exponential functions to estimate exponential growth; for example, it holds in the case of *linear bias* where consumers perceive exponential growth using $f(i,t,\theta) = 1 + it$. Linear bias is an approximation that many consumers appear to use when intuitively extrapolating exponential growth (Eisenstein and Hoch (2005)).

C. EG Bias and Perceived Returns to Saving at Different Horizons

PROPOSITION 1: Let $FV(i,t) = PV \cdot (1+i)^t$ be the actual future value of savings given a present value PV, interest rate *i*, and time horizon *t.* Let $FV^p(i,t,\theta) = PV \cdot f(i,t,\theta)$ be the future value perceived by a consumer with EG bias, where $f(i,t,\theta)$ satisfies properties (a)-(f) above. Then:

1. *A consumer with EG bias underestimates the actual future value, i.e. for any* $\theta > 0$, $FV(i,t) - FV^p(i,t,\theta) > 0$.
2. *Greater EG bias implies more severe underestimation, i.e.* $FV(i,t) - FV^p(i,t,\theta)$ *is strictly increasing in* θ.
3. *For any degree of EG bias* $\theta > 0$, *underestimation is greater in level terms over longer time horizons, i.e. for* $t' > t$, $FV(i,t') - FV^p(i,t',\theta) > FV(i,t) - FV^p(i,t,\theta)$.

Proof:

1. $FV(i,t) = PV \cdot (1+i)^t > PV \cdot f(i,t,\theta) = FV^p(i,t,\theta)$, where the first equality holds by property (a), the inequality holds by property (c), and the last equality holds by definition.
2. $FV(i,t) - FV^p(i,t,\theta)$ is strictly increasing in θ because $FV(i,t)$ is independent of θ by definition, and $FV^p(i,t,\theta)$ is decreasing in θ by property (c).
3. Underestimation is more severe at longer time horizons by property (f). This follows directly from substituting $FV^p(i,t,\theta) = PV \cdot f(i,t,\theta)$ and

$$FV(i,t) = FV^{p}(i,t,0) = PV \cdot (1+i)^{t} \ .$$

D. Exponential Growth Bias and Payment/Interest Bias

Establishing a relationship between EG bias and payment/interest bias is a bit more involved. Recall that the formula relating the interest rate to a loan payment m, maturity t, and loan principal amount L is

$$m = Li + \frac{Li}{\left[(1+i)^{t} - 1\right]}$$

Let $i^{*} = i^{*}(m,L,t)$ be the true interest rate, defined as the implicit solution for i in the equation:

$$G(m,L,i,t) = m - Li - \frac{Li}{\left[(1+i)^{t} - 1\right]} = 0$$

Define the perceived rate for a consumer with EG bias as i^{p}, where the perceived rate is the solution for *i* in:

$$\widetilde{G}(m,L,i,t,\theta) = m - Li - \frac{Li}{\left[f(i,t,\theta) - 1\right]} = 0$$

To be as conservative as possible, if this equality has multiple solutions we define the perceived rate as the highest solution:[3]

$$i^{p}(m,L,t,\theta) \equiv \sup\left\{ i \mid \widetilde{G}(m,L,i,t,\theta) \geq 0 \right\}$$

PROPOSITION 2: Suppose that $mt > L$ (guaranteeing that a positive solution for the actual interest rate exists). Consider a consumer with EG bias; i.e., someone who estimates exponential growth by $f(i,t,\theta)$ satisfying properties (a)-(f).

1. *The consumer with EG bias has payment/interest bias and strictly underestimates the true interest rate; i.e., for* $\theta > 0$, $i^{p}(m,L,t,\theta) < i^{*}(m,L,t)$.
2. *The degree of payment/interest bias is always weakly more severe the higher is* θ, *and is also strictly more severe for any positive perceived interest rate; i.e., if* $i^{p}(m,L,t,\theta') > 0$, *then for any* $\theta'' > \theta'$, $i^{p}(m,L,t,\theta'') < i^{p}(m,L,t,\theta')$.
3. *Payment/interest bias is more severe at short maturities under general conditions. Define pairs* $\left[(m',t'),(m'',t'')\right]$ *as payment/maturity combinations with an equal true APR* i^{*}, *i.e. for which (holding L constant):*

$$i^{*}(m',L,t') = i^{p}\left(m',L,t',0\right) = i^{*}(m'',L,t'') = i^{p}\left(m'',L,t'',0\right)$$

---

[3] Since we are interested in proving that EG bias leads to underestimation of the true rate, we choose the highest perceived rate that solves the problem.

Then if $t' < t''$ *and* $\theta > 0$ *, and with a functional form assumption (detailed below) on* $f(i,t,\theta)$ *, then*

$$i' = i^p\left(m', L, t', \theta\right) < i'' = i^p\left(m'', L, t'', \theta\right)$$

Proof: First we establish that $i \in [0, m/L]$. By l'Hopital's Rule,

$$\lim_{i \downarrow 0} G(m,L,i,t) = m - \frac{\partial(Li)/\partial i}{\partial[(1+i)^t - 1]/\partial i}$$
$$= m - \frac{L}{t}$$

The maintained hypothesis $mt > L$ thus implies $\lim_{i \downarrow 0} G(m,L,i,t) > 0$. For $i = m/L$

$G(m,L,m/L,t) = \dfrac{-m}{(1+m/L)^t - 1} < 0$, and $G(m,L,i,t) < 0$ for all $i > m/L$ as well. Hence

$$\{i \mid G(m,L,i,t) \geq 0\} \subseteq [0, m/L],$$

implying that $i^*(m,L,t) \in [0, m/L]$.

1. Arguments along the lines of Milgrom and Roberts (1994) establish part 1 of the proposition. Now $\tilde{G}(m,L,i,t,\theta) < \tilde{G}(m,L,i,t,0) = G(m,L,i,t)$. The first step holds because $\tilde{G}(m,L,i,t,\theta)$ is strictly increasing in $f(i,t,\theta)$, and $f(i,t,\theta)$ is strictly decreasing in θ. Hence,
$$\left\{i \mid \tilde{G}(m,L,i,t,\theta) \geq 0\right\} \subset \{i \mid G(m,L,i,t) \geq 0\}$$
which implies that
$$i^p(\theta) = \sup\left\{i \mid \tilde{G}(m,L,i,t,\theta) \geq 0\right\} < i^* = \sup\{i \mid G(m,L,i,t) \geq 0\}.$$
2. Similar arguments establish part 2 of the proposition. $\tilde{G}(m,L,i,t,\theta)$ is strictly decreasing in θ, implying that for $\theta'' > \theta'$,
$$i^p(\theta'') = \sup\left\{i \mid \tilde{G}(m,L,i,t,\theta'') \geq 0\right\} < i^p(\theta') = \sup\left\{i \mid \tilde{G}(m,L,i,t,\theta') \geq 0\right\}.$$

Remarks: In order to apply theorem #1 in Milgrom and Roberts (1994), we need to establish an interval on the interest rate for which the function G takes on both a positive and negative value (thereby establishing that it takes on a value of zero within the interval, meaning that a positive solution for the actual interest rate exists). Maintaining the assumption $mt \geq L$ assures this, and bounds the interval for the solution at $0 \leq i^* \leq m/L$. Intuitively, the assumption $mt \geq L$ states that the sum total of the payments must be greater than the loan principal, and it leads to an interval $0 \le i^* \le m/L$ in which interest rate must eventually repay the loan (rewriting, the condition is $m \ge Li$, which states that the periodic payment must at least cover the per-period interest charges).

3. First, define $i' = i^p(m', L, t', \theta) = \sup\{i \mid \tilde{G}(m', L, i, t', \theta) \ge 0\}$ and $i'' = i^p(m'', L, t'', \theta) = \sup\{i \mid \tilde{G}(m'', L, i, t'', \theta) \ge 0\}$. Note that it is sufficient to show that $\tilde{G}(m'', L, i', t'', \theta) > 0$ in order to show that $i' = i^p(m', L, t', \theta) < i'' = i^p(m'', L, t'', \theta)$. Rewriting,

$$\tilde{G}(m'', L, i', t'', \theta) = m'' f(i', t''\theta) - m'' - Li' f(i', t'', \theta)$$

And therefore the inequality that must hold is $m'' f(i', t'', \theta) - m'' - Li' f(i', t'', \theta) > 0$. Intuitively, this states that if the function $\tilde{G}(m'', L, i, t'', \theta)$ is positively evaluated at the interest rate that solves the problem $\tilde{G}(m', L, i, t', \theta) = 0$, then the solution solving $\tilde{G}(m'', L, i, t'', \theta) = 0$ must be a higher perceived rate.

Note also that by the definition of m'',

$$\tilde{G}(m'', L, i^*, t'', 0) = m'' f(i^*, t'', 0) - m'' - Li^* f(i^*, t'', 0) = 0$$

Rewriting,

$$m'' = \frac{Li^* f(i^*, t'', 0)}{f(i^*, t'', 0) - 1},$$

And substituting into $m'' f(i', t'', \theta) - m'' - Li' f(i', t'', \theta) > 0$ we have

$$\frac{Li^* f(i^*, t'', 0)}{f(i^*, t'', 0) - 1} f(i', t'', \theta) - \frac{Li^* f(i^*, t'', 0)}{f(i^*, t'', 0) - 1} - Li' f(i', t'', \theta) > 0 .$$

Multiplying through by $f(i^*, t'', 0) - 1$ gives

$$Li^* f(i^*, t'', 0) f(i', t'', \theta) - Li^* f(i^*, t'', 0) - Li' f(i', t'', \theta) f(i^*, t'', 0) + Li' f(i', t'', \theta) > 0 .$$

The loan principal cancels and the condition becomes:

$$i^* f(i^*, t'', 0) f(i', t'', \theta) - i^* f(i^*, t'', 0) - i' f(i', t'', \theta) f(i^*, t'', 0) + i' f(i', t'', \theta) > 0 .$$

Dividing through by i' gives:

$$\frac{i^*}{i'} f(i^*, t'', 0) f(i', t'', \theta) - \frac{i^*}{i'} f(i^*, t'', 0) - f(i', t'', \theta) f(i^*, t'', 0) + f(i', t'', \theta) > 0$$

The ratio $\frac{i^*}{i'}$ can be rewritten by noting that because

$$\tilde{G}(m', L, i', t', \theta) = \tilde{G}(m', L, i^*, t', 0) = 0 ,$$

$$m' = \frac{Li^* f(i^*, t', 0)}{f(i^*, t', 0) - 1} = \frac{Li' f(i', t', \theta)}{f(i', t', \theta) - 1}$$

And therefore,

$$\frac{i^*}{i'} = \frac{f(i^*, t', 0) f(i', t', \theta) - f(i', t', \theta)}{f(i^*, t', 0) f(i', t', \theta) - f(i^*, t', 0)}$$

Returning to the inequality above we see that it becomes:

$$\begin{aligned} & f(i', t'', \theta) f(i^*, t', 0) f(i^*, t'', 0) + f(i', t', \theta) f(i', t'', \theta) f(i^*, t', 0) + f(i', t', \theta) f(i^*, t'', 0) \\ & - f(i', t', \theta) f(i', t'', \theta) f(i^*, t''.0) - f(i', t', \theta) f(i^*, t', 0) f(i^*, t'', 0) - f(i', t'', \theta) f(i^*, t', 0) > 0 \end{aligned}$$

This expression states the inequality solely in terms of $f(i, t, \theta)$, and can be further simplified by noting that $f(i^*, t, 0) = (1 + i^*)^t$. Whether it holds in a particular instance depends on the functional form of $f(i, t, \theta)$. We have confirmed that it holds for the general parameterization of EG bias that we use in the text above.

Appendix C. Data Construction

This appendix contains details on each of the variables included in our empirical tests. Our sample frame for these tests is the 4,103 households in the 1983 *Survey of Consumer Finance*’s “cleaned area probability sample” and “high income sample” (see variable b3001). We drop the 159 “area probability excluded observations” that have incomplete data. Unless otherwise noted the SCF variables have no missing values due to perfect response or imputation.

RHS variables: Payment/Interest Bias, and Controls

Variables	Definitions based on SCF variable(s)
Payment/interest bias	*bias* = [*perceived rate – actual rate*]: *actual rate* (b5521) is an APR constructed by the SCF (and validated by us) based on the respondent’s self-supplied *repayment total* response to question b5516 or b5517. *perceived rate* is constructed from b5518 and b5519; these questions ask the respondent to impute an interest rate based on the respondent’s *repayment total.* *bias unknown* = 1 unless both *perceived rate* and *repayment total* are supplied. *addon* = 1 if (*perceived rate* = *add-on rate)*: *add-on rate* is the actual simple interest rate associated with the *repayment total*. This rate does not account for the declining balance implied by the survey’s scenario; e.g., the *add-on rate* on a *repayment total* of $1,200 is 20%, while the APR is 35%.
Male head of household	b3126=1.
Age of head	From b4503.
Race of head	From b3111.
Education category of head	From b3113, counting those who have junior college as highest attainment (b4507=1 & b4505<16) as “some college.”
Risk aversion/attitude (“financial risks” categories)	Non-missing categories constructed directly from b5403; 94 nonresponses grouped into one category.
Patience/liquidity attitude (“tie up money” categories)	Non-missing categories constructed directly from b5404; 114 nonresponses grouped into one category.
Borrowing attitude (“thinks buying on credit” categories)	Non-missing categories constructed directly from b5501; 37 nonresponses grouped into one category.
Expects to receive an inheritance	Categorical variable based on b4551: does not expect/expects/nonresponse.
Expected retirement age	Categorical variable based on b4519: < 50, 50, 51-54, 55, 56-59, 60, 61, 62,

	63-64, 65, 66-69, 70, >70, never, never worked full-time, already retired, nonresponse.
Expected years before leave current job	Categorical variable based on b4551: one category for each year, top-coded at 11, with separate categories for "never," no current job, and nonresponse.
Marital status	Based on b3112; binary variable that=1 if household head is married or lives with partner.
Household size	Categories from b3101, top-coded at 7.
Employment status	Head works full time: binary variable based on b4511, counting category 1 as full time. Spouse/partner works: Binary variable based on b4611, counting categories 1, 2, and 3 as working (includes those laid off who expect to return).
Health	Self-reported health status: excellent/good/fair/poor; we take categories for head (spouse/partner) directly from b4509 (b4609).
Homeownership	Binary variable based on b3702: we count category 1 as homeowners.
Industry category	14 Categories taken directly from b4539 (Census/CPS major industry group); observations with missing values are dropped.
Occupation category	For head; 8 categories taken directly from b3114 (self-employment category in subsumed in broader self-employment definition directly below).
Self-employed	Binary variable set to 1 if any of the following hold: - Head lists occupation as self-employed (b3114) - Head lists self as employer (b4540) - Household reports nonzero business income (b3206, b3512) - Household has ownership and management interest in a business (nonzero b3502).
Household wage category	Percentiles, constructed from b3205. Categories are constructed to have roughly equal frequencies, but must allow for the mass point at zero in the lowest category (1). To adjust for this, categories 2, 9, and 10 have relatively small frequencies.
Years in current job	From b4543: 10 categories, one for each year, top-coded at 10 years.
Pension coverage	From b4512; 1 = head's job provides pension and/or thrift benefits.
Social Security + pension wealth	11 categories (including one for missing values), constructed from b3317. Nonzero categories have roughly equal frequencies.

Uses advice on saving and investment decisions	Categorical variable constructed from b5340-b5347, which asks respondent “whether he/she sought advice concerning savings and investment decisions” from different sources. We categorize as: no advice/friends and family only/professional (accountant, banker, stockbroker, tax advisor, lawyer, financial advisor, insurance agent)/other.
Uses ATM	b5301 (ATM uses per “typical year”)>0.
Denied//discouraged/turned down for credit	Binary variable = 1 if: household was turned down for credit, or did not get as much credit as it wanted, “in the past few years” (b5522), AND did not end up obtaining the desired credit (b5525), OR: household had, “in the past few years… thought about applying for credit… but changed their mind because… might be turned down” (b5526); missing values dropped.
Owns a credit card	Binary variable = 1 if household has a bank card (b4108>0) or store card (b4114+b4117>0).
Compares loans terms on price or non-price margins	Binary variable = 1 if respondent reported that “size of the loan,” “size of the monthly payments,” “security for collateral for the loan,” or “size of the down payment” would be “the most important… if you were going to use credit to purchase a car” (b5513).
Net worth categories (added to some of the stockholding specifications)	Net worth excluding pension and Social Security wealth (b3323), categorized into deciles except that we impose the restriction that the bottom decile include only those with negative and zero net worth; this makes the bottom “decile” somewhat smaller, and the next decile somewhat larger, than the top 8 deciles.
Large recent purchase characteristics:	Nonzero only if household “purchased a vehicle, large item for the home, a recreation item, or home improvements, that cost $500 or more within the previous year:” b5601=1.
Purchase month and year	Binary variables for month*year constructed from b5603 and b5604.
Purchase: cost	log(b5605), replaced with zero if no purchase.
Purchase purpose	We constructed 14 categories of purchases from the more disaggregated b5602; includes category for no purchase.

LHS variables: Outcomes of Interest

Financed a large recent purchase using non-mortgage installment debt	Binary variable = 1 if: - Household made large recent purchase (b5601=1, see above) - Installment loan used (b5606=11 or b5606=12).
Net worth	b3323, which excludes pensions and Social Security. In constructing our net worth percentile variable we account for the small mass point at zero.
Short-term installment debt	Total amount outstanding on non-mortgage loans with regular payments (b4202). In our regressions we scale this by household income (b3201). Total income is <=0 in only two cases in our base sample of 3,911 observations, and never <=0 in our analysis samples for debt/income, where we restrict the sample to those with nonzero debt.
Long-term debt	(total debt outstanding – short-term installment debt) = (b3320 – b4202).
Owns any stock	Binary variable = 1 if households owns any publicly traded stock or non-money market mutual funds: b3462>0.
Stock share of financial assets	(b3462)/(financial assets), where: Financial assets = b3302 = demand deposits+money market+bonds+stocks+mutual funds+trust accounts.
Stock share of total assets	B3462/total assets, where we define total assets as financial assets + home value (b3708).
Certificate of Deposit (CD) ownership	Total dollar amount from b3453; any/share of financial assets/share of total assets defined as for stocks.
Saving rate category in 1982	Categories directly from b5406 as listed in Table 7; observations with missing values dropped from specifications in Table 7.
Uses any advice	Binary variable constructed from b5340-b5347, which asks respondent “whether he/she sought advice concerning savings and investment decisions” from different sources: accountant, banker, stock broker, tax advisor, lawyer, spouse, friend or relative, financial advisor, media, insurance agent, employer, other source. We set the variable = 1 if respondent reports using any of these sources.
Uses professional advice	= 1 if respondent reports using advice from any of: accountant, banker, stockbroker, tax advisor, lawyer, financial advisor, insurance agent.
Bond owner	= 1 if b3458>0.

Own company stock share	b3466/b3462, defined only for b3462>0.
Number of stocks owned	b3468, defined only for b3462>0. Public equities only, does not include mutual funds.
Number of trades	b3469, only counts public equity trades using a broker (SCF sets to zero otherwise).

References

Arnott, David R., 2006, Cognitive Biases and Decision Support Systems Development: A Design Science Approach, *Information Systems Journal* 16, 55-78.

Arnott, David R., and Peter O'Donnell, 1997, An Experimental Study of the Impact of a Simple Computer-based Decision Aid on the Forecast of Exponential Data, Working Paper.

Ebersbach, Mijam, and Friedrich Wilkening, 2007, Children's Intuitive Mathematics: The Development of Knowledge about Nonlinear Growth, *Child Development* 78, 296-308.

Eisenstein, Eric, and Stephen Hoch, 2005, Intuitive Compounding: Framing, Temporal Perspective, and Expertise, Working Paper.

Jones, Gregory, 1984, Perceptions of Inflation: Polynomial not Exponential, *Perception and Psychophysics* 36, 485-487.

Kemp, Simon, 1984, Perceptions of Changes in the Cost of Living, *Journal of Economic Psychology* 5, 313-323.

Keren, G., 1983, Cultural Differences in the Misperception of Exponential Growth, *Perception and Psychophysics* 34, 289-293.

Milgrom, Paul, and John Roberts, 1994, Comparing Equilibria, *American Economic Review* 84, 441-59.

Wagenaar, William A., and Sabat Sagaria, 1975, Misperception of Exponential Growth, *Perception and Psychophysics* 18, 416-22.

Wagenaar, William A., and H. Timmers, 1978, Extrapolation of Exponential Time Series is not Enhanced by Having More Data Points, *Perception and Psychophysics* 24, 182-184.

Wagenaar, William A., and H. Timmers, 1979, The Pond-and-Duckweed Problem: Three Experiments on the Misperception of Exponential Growth, *Acta Psychologica* 43, 239-251.