

Survey Evidence on Habit Formation

Jiannan Zhou*

November 4, 2019
(Click here for the latest version)

Abstract

Habit formation captures the diminishing of human response to repetitive stimulation.¹ Models with habit formation have explained many important economic phenomena, such as the equity premium puzzle. The literature, however, disagrees or is uninformative on its micro evidence. Does habit formation exist? How fast does habit change? Are additive and multiplicative habits consistent with people's behavior? How strong is habit formation relative to keeping up with the Joneses? Can habit formation explain the Easterlin paradox? To answer these five questions, I design and field a survey eliciting ten preference parameters on habit formation. I find that both internal and external habits exist, with the latter accounting for about 17% of habit. Adjustment and cognition costs do not explain habit formation. Habit depreciates about two-thirds per year. The evidence rejects both additive and multiplicative habits. Habit formation is about as strong as keeping up with the Joneses. Combining habit formation with keeping up with the Joneses could potentially explain the Easterlin paradox.

JEL: E21, G12, D60, I31.

Keywords: habit formation, micro evidence, keeping up with the Joneses, Easterlin paradox.

*Department of Economics at the University of Colorado Boulder. jiannan.zhou@colorado.edu.
jiannanzhou.com. I am indebted to professors Alessandro Peri, Xingtan Zhang, Martin Boileau, and especially Miles Kimball for their advice and support. I thank professors Richard Mansfield, Carlos Martins-Filho, Adam McCloskey, Sanjai Bhagat, Scott Savage, Terra McKinnish, Tony Cookson, Nathalie Moyen, Scott Schuh, Jin-Hyuk Kim, Nicholas Flores, Charles de Bartolomé, Shuang Zhang, and numerous seminar and conference participants for helpful comments.

¹This notion of habit formation is what the current habit formation models in macroeconomics and finance are trying to capture. It is different from the day-to-day notion of the cue-routine-reward habit. As it is already a received label in the literature, I will continue referring to this phenomenon as habit formation. This definition of habit formation sets it apart from path dependence, which is sometimes confused with habit formation. It is also worth noting that habit formation is different from desensitization, which would imply reduced response to small changes. Habit formation increases response to small changes.

1 Introduction

Habit formation² has helped explain important phenomena in, among others, asset pricing, business cycles, and economic growth,³ and is a fixture in macroeconomic models for policy analysis.⁴ Yet, literature is not certain of its micro evidence (Cochrane, 2017). Micro evidence is essential in transforming “interesting calculations and properties” of models to credible predictions (Hansen and Heckman, 1996). This paper provides a comprehensive set of micro evidence on habit formation, as follows.

First, habit formation exists, as a phenomenon distinct from adjustment and cognition costs. Depending on the source of habit, models with habit formation can be categorized into internal habit (habit formed on self past behavior) and external habit (habit formed on other people’s past behavior). Literature investigating evidence of habit formation has mostly focused on the existence of internal habit formation and drawn mixed conclusions (see column (1) of Table 1). Studies based on macrodata tend to find significant evidence of internal habit formation, while studies based on microdata find it to be relatively less significant. Existence of external habit formation is much less studied (column (2) of Table 1) while a lot of theoretical papers assuming it (see, e.g., Abel, 1990; Campbell and Cochrane, 1999; Smets and Wouters, 2007; Dou et al., 2017). Without a consensus on the existence of habit formation, literature has started investigating whether habit formation is capturing other phenomena. Flavin and Nakagawa (2008) propose adjustment costs as a deeper explanation for habit formation. Matyskova et al. (2019) argue habit as a way to economize cognition costs. In this paper, I provide evidence on the existence of both internal and external habit formation and evidence that habit formation cannot be explained by adjustment and cognition costs.⁵ Few papers study the composition of internal and external habit formation (column (3) of Table 1). Allowing habit to be formed both internally and externally à la Grishchenko (2010), I estimate that external habit accounts for a small portion (about 17%) of habit.

Second, habit depreciates about two-thirds per year. Most specifications of habit

²Throughout this paper, habit refers to (real) total consumption habit. There is deep-habit literature that assumes people form habit on individual varieties of consumption, which is not the focus of this project.

³E.g., equity premium puzzle and stock market behavior (Constantinides, 1990; Campbell and Cochrane, 1999); excess smoothness and excess sensitivity of consumption (Fuhrer, 2000); causal effect of high growth on high saving (Carroll, Overland and Weil, 2000); simultaneous decline in investment and current account (Cardi and Müller, 2011).

⁴See, e.g., Smets and Wouters (2007); Christiano, Eichenbaum and Trabandt (2015); Dou et al. (2017).

⁵Biologists have found evidence on habit formation in humans and all animals, including single-celled amoeba that does not have a nervous system (Folger, 1926). My result implies that people form habits for reasons beyond adjustment and cognition costs. For example, Rayo and Becker (2007) argue from the evolutionary perspective that habituating living standards makes us more motivated to strive for more, which potentially helps with survival.

formation depend on two important parameters—habit depreciation rate⁶ and habit intensity. Literature has focused predominantly on the habit intensity parameter (Havranek, Rusnák and Sokolova, 2017; column (4) of Table 1) while treating the habit depreciation rate largely as a free parameter whose value is often chosen to improve model fit with data. Two potential reasons are the relative lack of evidence on this parameter and the massive disagreement between the few existing micro and macro estimates (column (5) of Table 1). I provide a micro estimate that is much closer to, though still relatively far away, from typical macro estimates. To illustrate the importance of this parameter, I show that simply changing the value of this parameter can significantly degrade the performance of some habit formation models.

Third, neither additive nor multiplicative habit is consistent with people’s behavior. Almost all current habit formation models in the literature are assuming either one of these two habit utility functions (column (6) of Table 1). The conclusions drawn from these studies are, therefore, joint tests and estimates with these specifications. In a general utility function naturally nesting these two formulations, I propose and implement four tests on the micro validity of the preference specifications. I find that both types of habit utility function fail the tests by large margins. Even though these two particular specifications are not consistent with the evidence, estimates of the signs of utility derivatives are consistent with habit formation,⁷ suggesting that other specifications of habit formation could be consistent with the micro evidence.⁸

Fourth, the effect of habit formation on utility is about as strong as that of keeping up with the Joneses. As two important interdependent preferences, keeping up with the Joneses allows interpersonal dependence, while habit formation allows intertemporal dependence (in internal and external habits) as well as interpersonal dependence (in external habit). Literature finds a strong effect of keeping up with the Joneses (Luttmer, 2005; Lewbel et al., 2018; De Giorgi, Frederiksen and Pistaferrri, 2019) but disagrees on the relative strength of these two phenomena. Ravina (2007) finds habit formation to be about 70% stronger than keeping up with the Joneses, while Alvarez-Cuadrado, Casado and Labeaga (2015) conclude that internal habit formation is as strong as keeping up with the Joneses. I investigate this question without specifying particular functional forms for utility function and the speed that habit depreciates.

Fifth, habit formation joining forces with keeping up with the Joneses could generate the income-happiness pattern of the Easterlin paradox. Easterlin (1973, 1974) highlights the tension between cross-sectional positive correlation and time-series zero correlation between happiness and income and proposes keeping up with the Joneses as an explanation for its effect on averaging happiness. As happiness data accumulates over time, the literature discovers that the zero time-series correlation tends to hold only in the long-run while the short-run correlation is generally positive (Steven-

⁶I focus on depreciation rate rather than catch-up rate because the latter varies under different normalizations of habit while the former is invariant to such normalizations.

⁷That is, $u_{CH} > 0$ and $u_{HH} < 0$.

⁸I explore this possibility in another paper.

son and Wolfers, 2008; Sacks, Stevenson and Wolfers, 2012; Easterlin, 2017).⁹ Habit formation has been proposed as a potential explanation to the original version of the paradox¹⁰ (Easterlin, 1995) and in particular, to the relatively newly discovered temporal heterogeneity of the correlation (Clark, Frijters and Shields, 2008; Clark, 2016). To the best of my knowledge, evidence on whether habit formation can actually fit in this shoe is absent from the literature. Utilizing my aforementioned comprehensive set of evidence on habit formation, I conduct a semi-structural simulation and find that when coupling with keeping up with the Joneses, habit formation is able to generate the observed patterns of income and happiness across all the dimensions: cross-section, short-run, and long-run.

Providing this extensive set of evidence faces two challenges. First, I need a framework flexible enough to nest all current habit formation models that are heterogeneous along many dimensions while remaining agnostic of the existence of habit formation. Second, jointly estimating all the preference parameters requires variations that, to the best of my knowledge, do not exist in reality. To tackle the first difficulty, I take a flexible habit formation model with assumptions of minimal interference with the evidence I am providing. To tackle the second difficulty, I employ the method of eliciting preference parameters using hypothetical choices. While having its own limitations (response errors and biases),¹¹ this method has the advantages of creating variations absent in reality and of avoiding potential systematic biases of incentivized choices (e.g., misperceptions and lack of information).¹² The application of this method in economics can be traced back to at least Thurstone (1931) and has been widely used in eliciting reduced-form preference parameters in, among others, environmental economics, experimental economics, and health economics (Johnston et al., 2017; Ameriks et al., 2019; Jha and Shayo, 2019). The use of this method in eliciting structural preference parameters started since at least Barsky et al. (1997) and is commonly used thereafter in the literature (see, e.g., Kapteyn and Teppa, 2003; Sahm, 2007; Kimball, Sahm and Shapiro, 2008, 2009; Benjamin et al., 2014; Kimball et al., 2015; Benjamin et al., 2019).

In the flexible habit formation model, I identify ten preference parameters¹³ that are instrumental to the above set of evidence on habit formation. Guided by the

⁹I am aware of the controversy around the existence of the paradox. My purpose here is not to defend the paradox, as I do not provide new evidence on happiness measures, but to see how potentially can habit formation (and keeping up with the Joneses) affect the relationship between income (or consumption) and happiness.

¹⁰Kimball and Willis (2006) propose as a potential explanation of the paradox that happiness is time-intensive. They also survey several other potential explanations. In this project, I focus on the effects of habit formation and keeping up with the Joneses.

¹¹See also Diamond and Hausman (1994).

¹²See also Manski (2000).

¹³They are habit depreciation rate, time discount rate, external habit mixture coefficient, all ratios of utility derivatives up to the second-order, and two quantities concerning the existence of internal and external habits.

identification, I design and field a survey eliciting these preference parameters. Using Hamiltonian Monte Carlo, I jointly estimate all the parameters.

I present the flexible habit formation model and the survey instrument in section 2. Then I summarize the data and the statistical model that will be applied to the data (section 3). Section 4 contains the elicitation and estimate of each preference parameter of interest and how the estimates inform us of habit formation. I conduct the simulation on the Easterlin paradox in section 5. I discuss and check the robustness of my results in section 6. Section 7 concludes.

2 Methodology

2.1 Habit Formation Model

The breadth of the questions this paper intends to answer requires the habit model to be as general as possible so that it nests all specifications of interest. Generality means as few assumptions as possible. Still, enough structure must be put in place so that the model is useful to the purposes of this paper.

The agent maximizes

$$\mathbb{E}_0 \int_0^\infty e^{-\rho t} u(C(t), H(t)) dt$$

where C is individual spending, H habit, and ρ time discount rate. Henceforth the time index will be omitted when there is no ambiguity in doing so. I maintain expected utility and exponential time discounting.¹⁴ As is ubiquitous in the literature, I require that the utility function to be analytic and to satisfy positive monotonicity of consumption ($u_C > 0$) and diminishing marginal utility of consumption ($u_{CC} < 0$). These assumptions aid identification and estimation without interfering with the evidence this paper is providing. For example, I leave open whether habit affects the utility and the sign and magnitude of any derivative of the utility function with respect to habit. I allow the respondent's utility to depend on other variables (e.g., labor), but since they will be kept constant in the survey, abstracting from explicitly listing them as the arguments of the utility function results in no loss of generality. In the discussion of some survey questions where I do vary things other than self spending and habit (e.g., other people's spending), I will explicitly show the additional variables in the utility function. For clarity, I will save the details for relevant contexts.

I assume habit evolves according to

$$\dot{H} = \theta(C - H)$$

where θ is the habit depreciate rate. This formulation has been in the literature since at least Houthakker and Taylor (1970) and is the most commonly used habit

¹⁴Nearly all current habit formation models make these two assumptions.

Table 1: Estimates of Habit Parameters in Selected Literature

Preference parameters in the table are from specializations of following habit formation model:

$$u(C, H) = \begin{cases} v(C - \alpha H) & \text{Additive Habit} \\ v(C/H^\alpha) & \text{Multiplicative Habit} \end{cases} \quad \text{s.t. } \dot{H} = \theta((1 - \omega)C + \omega C_{others} - H)$$

where α is habit intensity, θ habit depreciation rate, and ω external habit mixture coefficient.

Study	Internal Habit (1)	External Habit (2)	ω (3)	α (4)	θ^3 (5)	Additive or Multiplicative ⁵ (6)
<i>A. Microdata</i>						
Naik, Moore (1996)	Y	(N)	(0)	0.081***	(Y)	(A)
Dynan (2000)	N	(N)	(0)	-0.038	(Y)	(A)
Guariglia, Rossi (2002)	N	(N)	(0)	-0.272***	(Y)	(A)
Lupton (2002)	Y	(N)	(0)	0.225***	9.2%/Y*** 4	(A)
Kapteyn, Teppa (2003)	Y	(N)	(0)	0.777***		(M)
Rhee (2004)	Y, N ¹	(N)	(0)	0.6	(Y)	(A)
Ravina (2007)	Y	N	0.025 ²	0.503***	(Q)	(A, M)
Browning, Collado (2007)	Y, N ¹	(N)	(0)	0.01-0.14	(Q)	(A)
Alessie, Teppa (2010)	Y	(N)	(0)	0.211**	(Y)	(A)
Iwamoto (2013)	N	(N)	(0)	-0.3787***	(Y)	(A)
Khanal et al. (2018)	Y	(N)	(0)	0.545***	(Y)	(A)
: (≥ 10 studies)	:	:	:	:	:	:
<i>B. Macrodata</i>						
Ferson, Constantinides (1991)	Y	(N)	(0)	0.64-0.97***	(M, Q, Y)	(A)
Führer (2000)	Y	(N)	(0)	0.80***	99.9%/Q***	(M)
Stock, Wright (2000)	Y, N ¹	(N)	(0)	-	(M, Y)	(A)
Smets, Wouters (2003)	(N)	Y	(1)	0.573***	(Q)	(A)
Lubik, Schorfheide (2004)	Y	(N)	(0)	0.57***	(Q)	(M)
Christiano et al. (2005)	Y	(N)	(0)	0.65***	(Q)	(A)
Smets, Wouters (2007)	(N)	Y	(1)	0.71***	(Q)	(A)
Adolfson et al. (2007)	Y	(N)	(0)	0.650***	(Q)	(A)
Korniotis (2010)	N	Y	0.79 ²	0.33 ²	(Y)	(A)
Grishchenko (2010)	Y	N	0.000***	0.895 ²	70.7%/Q***	(A)
Altig et al. (2011)	Y	(N)	(1)	0.76***	(Q)	(A)
: (≥ 65 studies)	:	:	:	:	:	:

Notes: The studies are selected based on citation, publication year, and number of parameters estimated. Characters in parenthesis and italics are assumed parameter values of the studies. 1: depends on goods or time horizon; 2: implied estimates; 3: M/Q/Y—habit depreciates fully at the end of a month/quarter/year; 4: geometric habit evolution speed: 0.071 (0.007); 5: A/M—additive/multiplicative habit. Significance levels: ** 5%, *** 1%.

evolution equation in the literature. Slightly different formulations of the evolution are used. But the difference either is simply a rescaling of the unit of habit (e.g., Constantinides, 1990) or disappears in steady state (e.g., Campbell and Cochrane, 1999). For general habit evolutions that potentially are nonlinear, I show that they are observationally equivalent to this linear habit evolution under reasonable conditions.¹⁵ I choose this habit evolution also for its natural unit, the same as that of consumption. For example, if a person has been consuming \$5,000 per month for as long as she can remember, then her current habit is also \$5,000 per month.

To answer the motivating questions, I need to see if the utility function depends on habit. If habit exists, I also need to know the values of the preference parameters, θ , ρ , ratios of utility derivatives up to the second-order ($-\frac{u_H}{u_C}$, $\frac{Hu_{HH}}{u_H}$, $\frac{u_{CH}}{u_{HH}}$, and $\frac{u_{CC}}{u_{HH}}$), and, in the extended models,¹⁶ external habit mixture coefficient and strength of habit formation relative to keeping up with the Joneses.

2.2 Survey Overview

To elicit the preference parameters, I specify simple hypothetical situations that identify the preference parameters of interest while controlling for potential confounding factors and response errors and biases. I implement the hypothetical situations in a survey.

My survey starts with a preamble module that specifies the basic hypothetical situation and instructs the respondents on the format of the survey. Nine core modules follow to elicit the preference parameters of interest.

The hypothetical situation is designed to be as simple as possible while still allowing the elicitation of what I am interested in and avoiding potential confounding factors that plague traditional data. The key is not realism, but simplicity. If the situation is simple enough, though hypothetical, people will be able to understand it. In particular, I want to free respondents from worrying about changes to the purchasing power of their money, about durable goods, and about changes in preferences. The exact wording of the hypothetical situation in the survey is the following.

Please answer all survey questions under the following hypothetical situation:

- There is no inflation and prices of everything stay the same over time.
- You rent the durable goods you consume, including residence, furniture, car, etc.
- Things you want don't change over time.
- People not mentioned in questions always spend \$5,000 per month.

¹⁵See section B of the appendix for the proof. The basic intuition of the proof is that the nonlinearity can be “transported” to the utility function which is agnostic of how habit affects it.

¹⁶See sections 4.4 and 4.5.

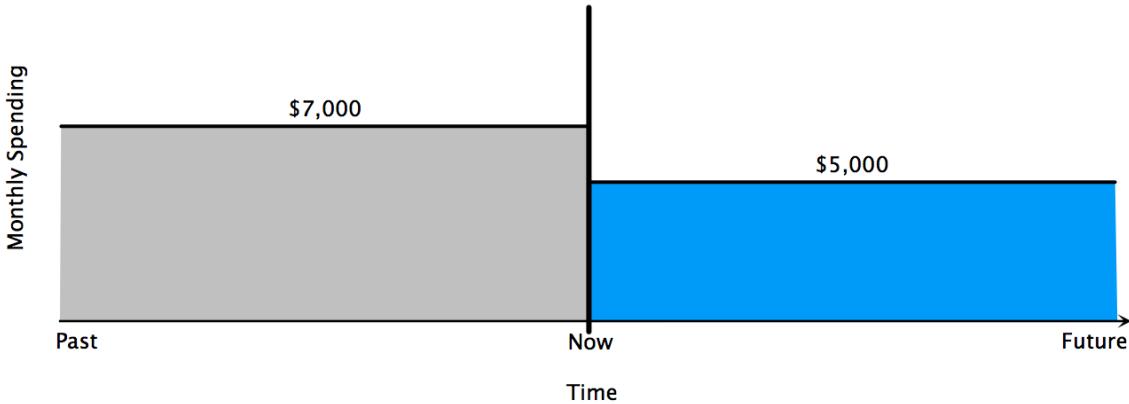


Figure 1: A Typical Monthly Spending Graph

- Everything else unmentioned in the questions is and stays the same.

I test respondents' understanding of this hypothetical situation. Only those who pass the test are able to proceed to the remaining modules in the survey.

The respondents did not know that they were answering a survey about habit formation. I told the respondents that the survey is about spending behavior. I do this for two reasons. First, to avoid potential confusion. More likely than not, a typical respondent does not know what the consumption habit is as we economists call it. Second, to avoid potential biases. I cannot prime them with habit formation while I want to test the very existence of it.

Because I am interested in how habit affects people's well being and habit responds to spending changes, I vary people's past and future spending profiles to elicit the preference parameters of interest. I create variations in both past and future spending paths because habit reveals itself in the way that the past affects the future. Since income affects utility through spending, I do not specify the income process except to tell the respondents that they can afford the spending profiles in the survey.

To make the representation of a spending profile intuitive and to simplify comparison across several of them, I draw a spending path in, as I call it, a monthly spending graph (Figure 1). Time is on the horizontal axis: past on the left, now in the middle, and future on the right. The bars on the graph represent monthly spending, which are drawn to scale and colored differently to help distinguish time horizons. In the example of Figure 1, the spending path represents spending \$7,000 per month in the past until now and \$5,000 per month in the future starting now. The respondents went through an instruction and a test on reading the monthly spending graphs before answering questions in the core modules.¹⁷

To alleviate the concern that in reality each person has only one past spending

¹⁷For details on this instruction, see section F.1 of the appendix.

path, I invoke the metaphor of parallel universes. In the parallel universes, everything is the same except for the spending profiles. I then ask the respondents which universe brings them a better future experience: how they feel in the future starting now.

Despite their advantages, surveys can be subject to response errors and biases. Nevertheless, it is always good to have more perspectives getting at the same question. This is especially valuable when traditional methods suffer from major limitations (e.g., alternative explanations) and papers using them disagree with each other. I build into the survey various details to minimize response errors and biases (see Figures 2 and 3 for screenshots of a typical survey question). For example, to reinforce the idea that the only variation between universes is in spending paths, I emphasize it at the start of every set of questions.¹⁸ To help the respondents compare different spending paths, I tell them in what time horizon the paths differ. To help them distinguish past experience from future experience, I allow them to express views on both experiences. I also reiterate definitions of the experiences of interest and highlight the key words—past and future—to further remind the respondents of which experience the question is about. To avoid respondents accidentally clicking an option different from the one they wanted to choose, I build the monthly spending graphs into the clickable choices. To inform them even further which choice they are choosing, I slightly darken the background of an option when their mouses hover over it and completely darken the background of the option of choice (Figure 4).

To make sure that my respondents are paying attention, I have a set of attention checks, ranging from explicit ones, like the quiz on the basic hypothetical situation at start of the survey, to implicit ones, such as time spent on the survey, quiz on the basic hypothetical situation at the end of the survey, and cross-wave consistency of their answers to the demographic questions. To encourage effort, I told the respondents about the existence of such attention checks but did not tell them where they are and what they look like in the survey. Also, to encourage more effort, I told them at the introduction section of the survey that a small (\$1) lottery with a winning chance of 1 to 100 will be rewarded to respondents of high-quality responses.¹⁹

To further minimize potential response errors and biases, I conducted two waves of the survey and changed the sequence of core modules and reversed the order of choices for questions in the second wave. I deal with remaining response errors and biases through a statistical model and robustness checks.²⁰

The survey responses do not directly tell me what values the preference parameters are, but give me information with which I can uncover the values of the preference parameters. In essence, the uncovering process is a mathematical proof that translates survey responses into values of the preference parameters of interest. Based on accuracy of the translation, I distinguish different orders of elicitation. For example,

¹⁸A set of questions corresponds to one or two preference parameters of interest and usually consists of 2 to 3 subquestions.

¹⁹Of the 515 responses I collected, 6 respondents were paid this small lottery.

²⁰See sections 3 and 6 for details.

- Imagine that Universes One and Two are identical except your monthly spending in the 'past'.
- Remember that past experience is how you felt about the 'past' until 'now'.

Which universe gave you a better **past** experience?

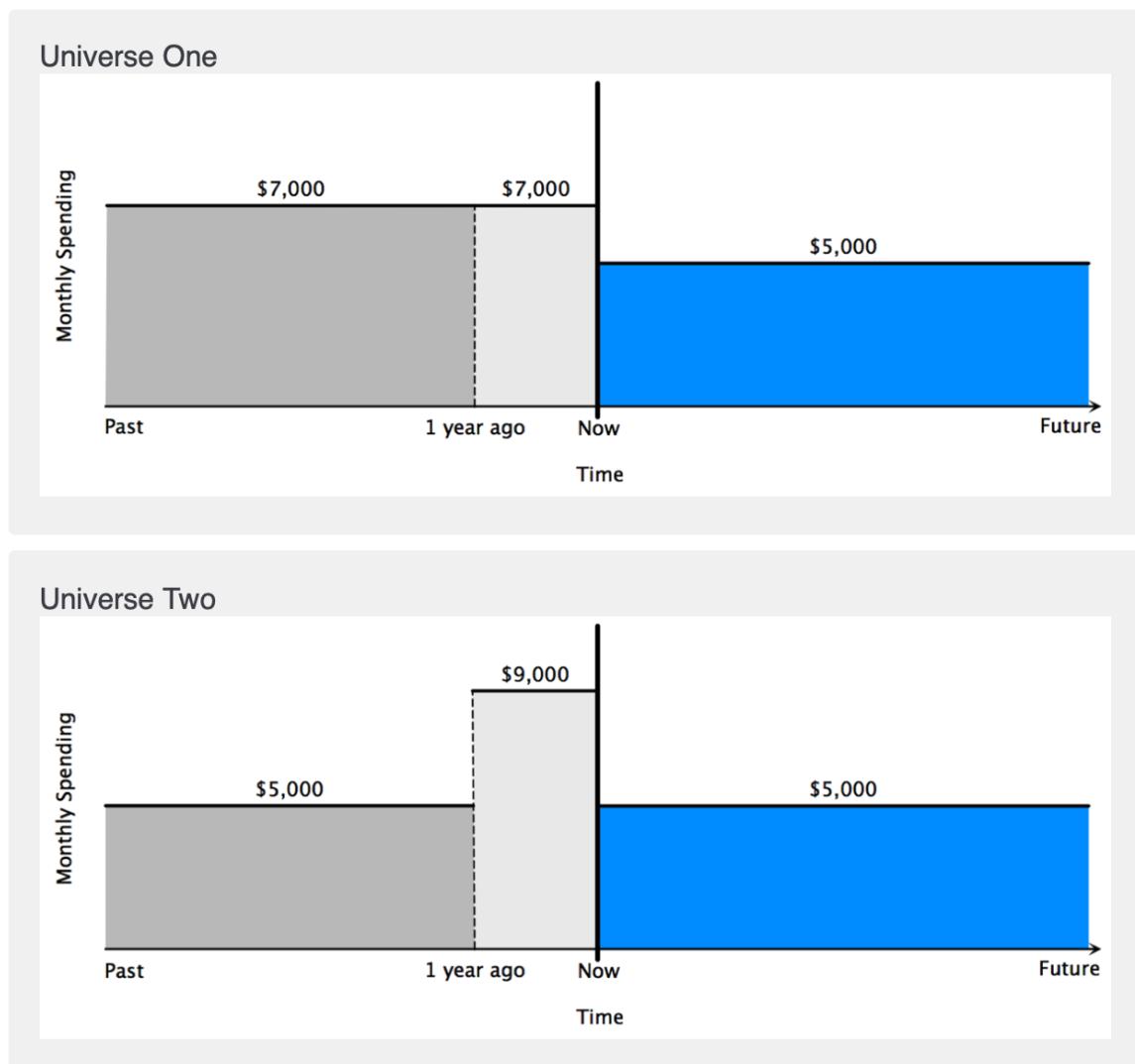


Figure 2: A Typical Survey Question (Part 1 of 2)

- Remember that future experience is how you feel about the 'future' starting 'now'.

Which universe will give you a better **future** experience?

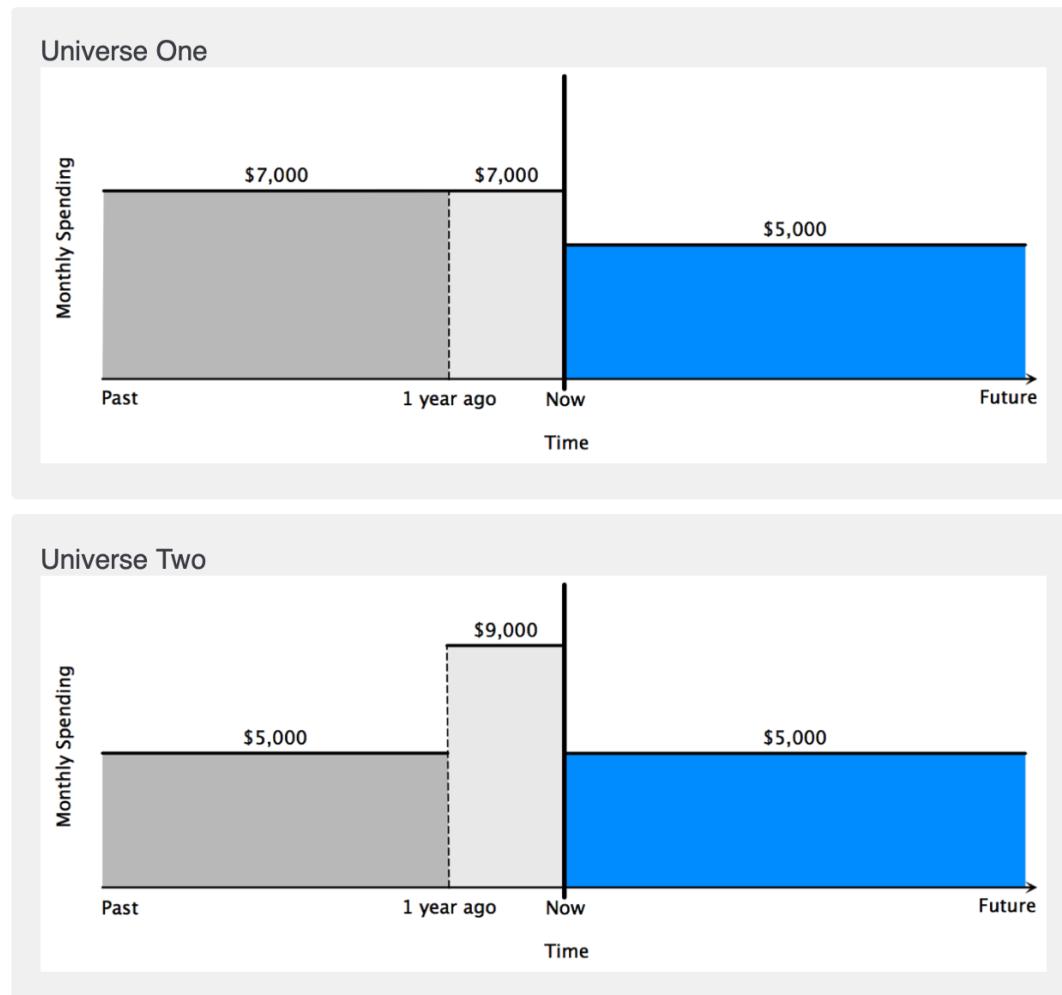
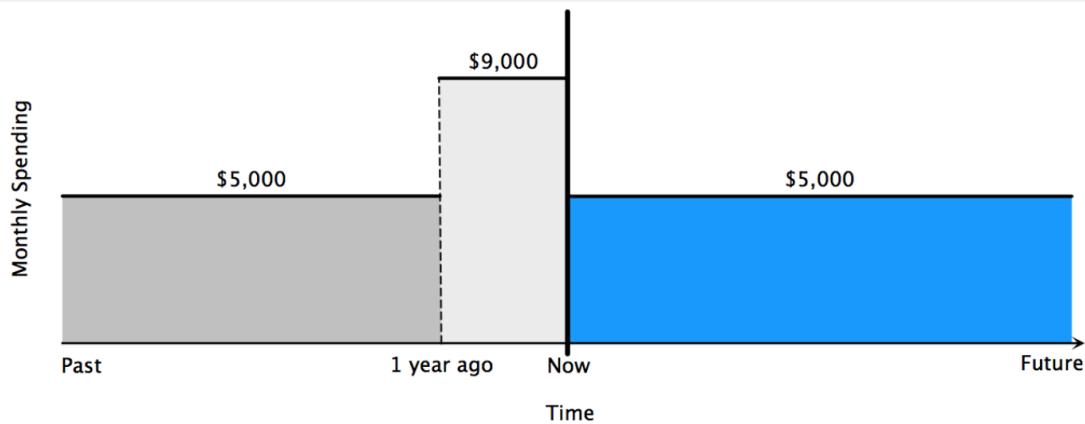
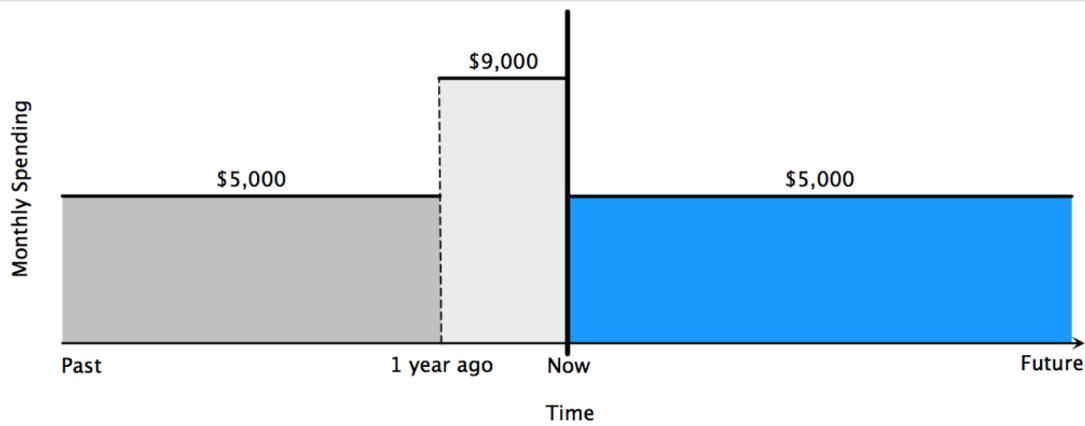


Figure 3: A Typical Survey Question (Part 2 of 2)

Universe One



Universe One



Universe One

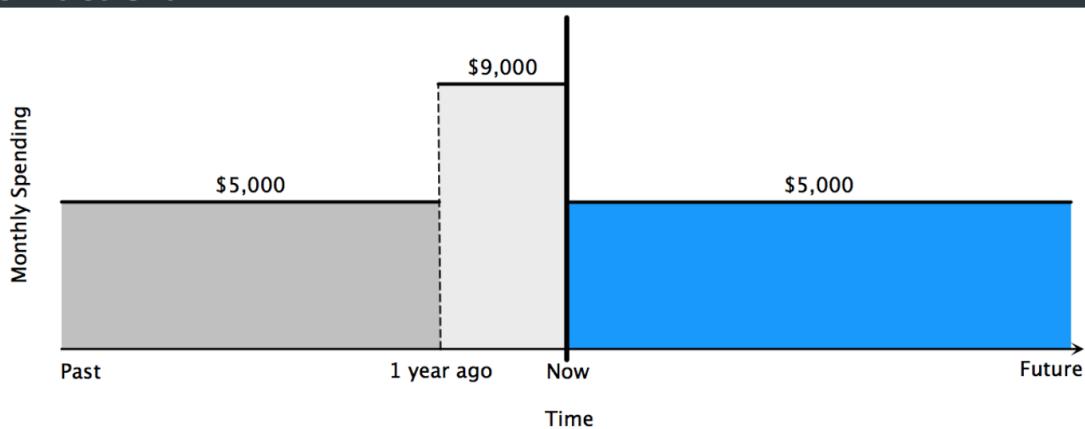


Figure 4: Option Appearances (Top to Bottom: Initial, Hover, Click)

first-order elicitation means that the translation error will be up to the remainder of a first-order Taylor expansion, while exact elicitation means that there will not be any translation error. In eliciting some of the preference parameters under the general habit model, translation error is unavoidable. However, my estimates show that even first-order elicited preference parameters can be very close to those estimated using traditional methods.²¹ The elicitation varies from one preference parameter to another.

3 Data and Statistical Model

3.1 Data

I conducted two waves of the survey on Amazon Mechanical Turk.²² Although this sample is less representative of the U.S. population than national probability samples such as Health and Retirement Study and Panel Study of Income Dynamics, it is more representative than lab samples (Berinsky, Huber and Lenz, 2012). Johnson and Ryan (2018) show that this sample can provide consistent and economically meaningful data. Many economics papers have utilized this sample (see, e.g., Oster, Shoulson and Dorsey, 2013; Kuziemko et al., 2015; Bordalo et al., 2016; Cavallo, 2017; Martínez-Marquina, Niederle and Vespa, 2019; Benjamin et al., 2019).

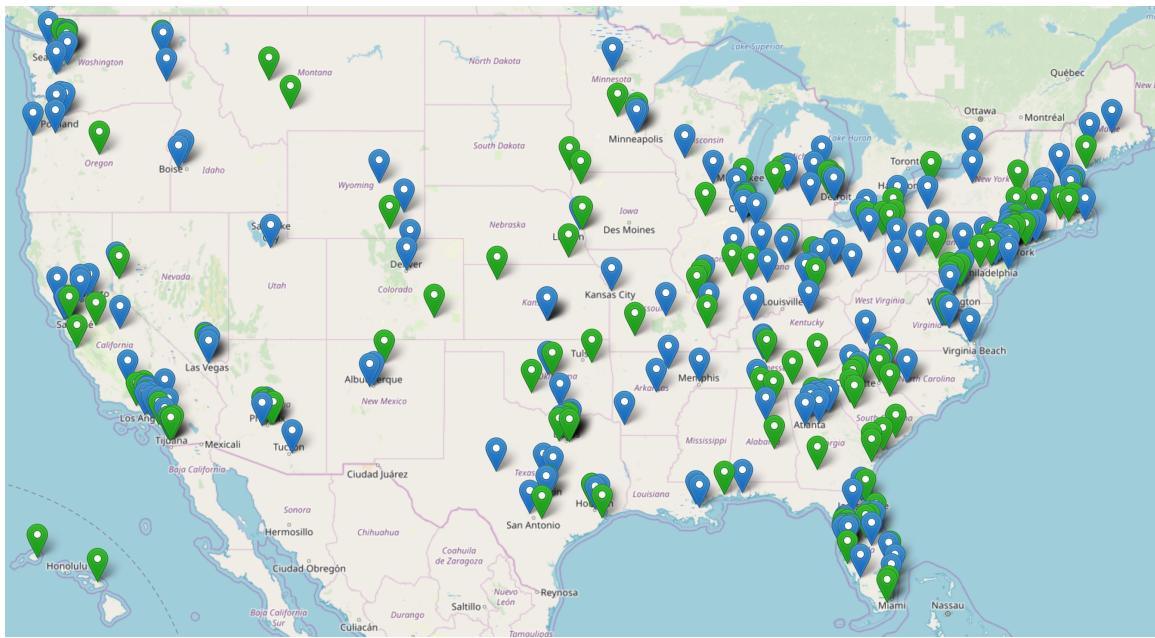
I restrict my respondents to U.S. residents. From the 301 first-wave respondents who expressed a willingness to participate in future studies, I randomly invited 200 respondents to participate in the second wave and got a response rate of about 75%. After excluding respondents located outside of the U.S., submitted duplicate responses, and sped through the survey, I am left with 348 and 140 valid responses from the two waves, respectively, by respondents across the U.S. (Figure 5).

In the survey, I collected demographic information on age, gender, household income, and household size. These sample characteristics are close to national counterparts (Table 2). At the time of the survey, a typical respondent is about 38 years old, lives with another person or two people in a household of an annual income of the range \$50,000 to \$60,000, is slightly more likely to be female if participates in only the first wave and slightly more likely to be male if participates in both waves, and spent less than half an hour on the survey.

I have nine core survey modules, each corresponding to one or two preference parameters of interest. Eight of the ten elicited preference parameters are identifiable to scale and estimated jointly for accuracy. As a result, response frequencies of individual parameters alone are not particularly informative, which are, therefore, relegated to the appendix (Table 25). The response frequencies of parameters identifiable only to sign are reported in section 4.

²¹See section 4.5 for details.

²²The first wave was conducted on July 23, 2018 and the second wave on August 11, 2018.



Notes: Green markers indicate locations of respondents participating in both waves while blue markers only in first wave.

Figure 5: Locations of Respondents

Table 2: Demographic Statistics

	First Wave	Second Wave	U.S.
Age, median	38	37	38
Household income, median	\$50,000 - \$60,000	\$50,000 - \$60,000	\$57,652
Female percentage	53.2%	47.9%	50.8%
Household size, mean	2.66	2.70	2.63
Time on survey, mean	28' 6"	25' 32"	

Notes: Household income is annual. The last column has the counterparts of the statistics at the U.S. national level. The sources of the U.S. statistics are 2018 Population Estimates (for age and female percentage), 2017 American Community Survey, and 2017 Puerto Rico Community Survey (for household income and size), all from U.S. Census Bureau.

3.2 Statistical Model

As discussed above, the survey design minimizes response errors and biases. In addition to getting rid of invalid responses from the sample, I deal with remaining response errors through the following statistical model. Potential response biases are dealt with through robustness checks (section 6)

I model an observed response for preference parameter x of individual i in wave w , $X_{i,w}$, as

$$X_{i,w} = \sum_k k \cdot 1(T_{k,\tilde{x}} \leq \tilde{x}_{i,w} \leq T_{k+1,\tilde{x}})$$

where the unobserved parameter value— $\tilde{x}_{i,w}$ —can be measured with error,

$$\tilde{x}_{i,w} = x_i + \varepsilon_{i,x,w}$$

, and $T_{\{k\},\tilde{x}}$ is the sequence of known thresholds informed by the elicitation of the parameter. The true parameter value for individual i , x_i , is drawn from $\mathcal{N}(\mu_x, \sigma_x^2)$. The individual-parameter-wave-specific response error $\varepsilon_{i,x,w}$ is drawn from $\mathcal{N}(0, \varsigma_{x,w}^2)$ independently of the true parameter value. In the robustness section, I allow the mean of the response errors to be non-zero and vary across waves. Estimates of the means of the more general response errors are indistinguishable from zero and that the estimates of the preference parameters are not significantly different from those under the zero-mean response error specification here. I assume that the parameters are drawn independently within a respondent. Because my respondents spread across the U.S. (Figure 5) and most likely did not know each other, I assume that responses are independent across respondents.

Allowing the response error to be persistent across waves (i.e., $Cov(\varepsilon_{i,x,1}, \varepsilon_{i,x,2}) = \sigma_{\varepsilon_x}^2 > 0$ and $\varsigma_{x,w}^2 = \sigma_{\varepsilon_x}^2 + \sigma_{\varepsilon_{x,w}}^2$), I arrive at the joint distribution of respondent i 's parameter x in the two waves of the survey

$$\begin{bmatrix} \tilde{x}_{i,1} \\ \tilde{x}_{i,2} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_x \\ \mu_x \end{bmatrix}, \begin{bmatrix} \sigma_x^2 + \sigma_{\varepsilon_x}^2 + \sigma_{\varepsilon_{x,1}}^2 & \sigma_x^2 + \sigma_{\varepsilon_x}^2 \\ \sigma_x^2 + \sigma_{\varepsilon_x}^2 & \sigma_x^2 + \sigma_{\varepsilon_x}^2 + \sigma_{\varepsilon_{x,2}}^2 \end{bmatrix}\right)$$

In section A of the appendix, I prove that the individuals' parameter values aggregate to the mean for the representative agent. That is,

$$x_R = \frac{1}{N} \sum_i x_i$$

where x_R is the value of the representative agent's parameter x . Given that almost all current habit formation models assume a representative agent, I will focus on the implication of my estimates for a representative agent model with habit formation. Since $x_R = \mu_x$, the estimate of interest is that of μ_x .

To increase accuracy, I estimate the parameters jointly. I use the Markov Chain Monte Carlo method Hamiltonian Monte Carlo that excels at high dimensional estimation. I choose uniform distributions as priors for the parameters, not only to

let data speak as much as possible but also establishes the equivalence between my maximum a posteriori (MAP hereafter) estimates and maximum likelihood estimates. I run ten Markov chains initialized from random diffuse starting points and collect 15000 iterations of warmup and 25000 draws of sampling. I report all three Bayesian point estimates²³ and the highest posterior density or mass interval (HPDI and HPMI henceforth).

4 Elicitation and Estimation

In this section, I show how the design of the survey and the estimates of preference parameters answer the questions on the micro evidence of habit formation. I break discussion of the existence of internal habit formation and external habit formation into two parts, for the latter involving model augmentation that is best done after discussing survey questions that do not require it.²⁴ First, let us look at existence of internal habit formation, habit depreciation speed, and additive versus multiplicative habits. Then, we turn to existence of external habit formation, composition of the two habits, and relative strength of habit formation and keeping up with the Joneses.

4.1 Existence of Internal Habit Formation

The fundamental characteristic of habit formation is the diminishing response to a repetitive stimulus. In the case of internal habit formation, the higher a person's past consumption (stimulation) is, the lower her future utility (response) is. As a measure of the intensity and persistence of past stimulation, habit increases with consumption. That is, habit formation is consistent with $Q_H \equiv u(C, H + \Delta h) - u(C, H) < 0$, but not with $Q_H \geq 0$, for $\Delta h > 0$.

To elicit the sign of Q_H , I vary the respondent's past spending profile so that the respondent's current habit varies across the parallel universes. I choose a common level of future spending for all the universes so that the only difference between the universes is the current level of habit when it comes to future experience. It is worth emphasizing again that I did not prime the respondents with habit formation in the survey, and that I do not make any assumption about the signs of derivatives of the utility function with respect to habit.

The survey question asks the respondents to pick the monthly spending graph that gives a better future experience (Figure 6). Preferring a universe with less past spending over one with higher past spending implies $Q_H < 0$. The responses to this question show that the average respondent chooses Universe One (Table 3), consistent with the existence of habit formation for the representative agent. My estimate confirms this (Table 4).

²³MAP together with posterior mean and median.

²⁴The later model augmentation does not affect results precede it. For more discussion on this, see section 4.4.

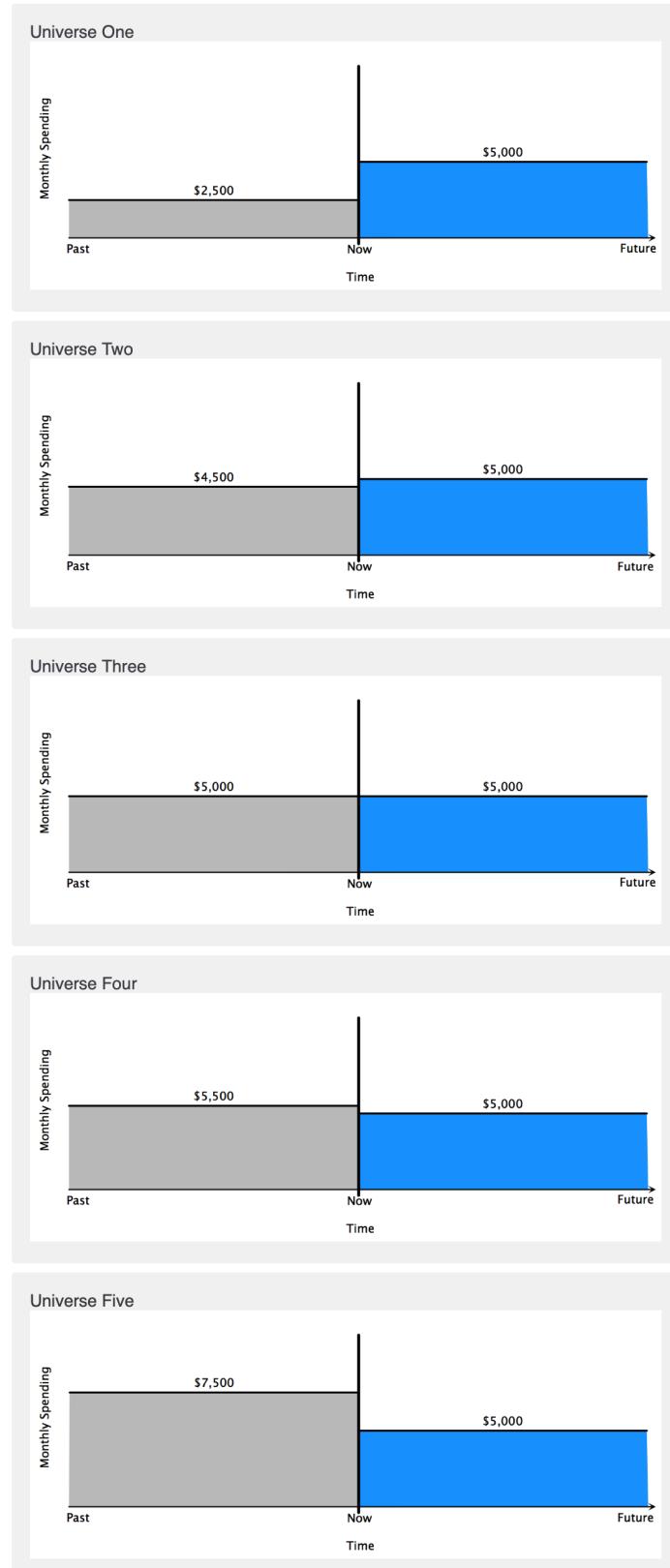


Figure 6: Monthly Spending Graphs: Existence of Habit Formation

Table 3: Response Frequency (Percentage) of Existence of Internal Habit Formation

Universe	1	2	3	4	5
1st Wave	56	3	9	1	30
2nd Wave	60	2	6	1	30

Table 4: Estimate of Sign of Q_H

	MAP	Mean	Median	99% HPMI
$\text{sgn}(Q_H)$	-1.00	-1.00	-1.00	[-1.00, -1.00]

By monotonicity of the utility function,²⁵ $Q_H < 0$ implies $u_H < 0$.

This survey question also enables me to distinguish between adjustment costs and internal habit formation. Flavin and Nakagawa (2008) argue that the adjustment cost of consumption commitments match data better and raise the possibility that adjustment cost is the structural explanation of habit formation. Chetty and Szeidl (2016) show that the difference between the two models matters for welfare and policy because consumption commitment can be abandoned quickly while habit cannot. To distinguish the two phenomena, imagine that only adjustment cost exists. As is currently modeled (Chetty and Szeidl, 2016), one pays adjustment cost whenever she changes her level of consumption. It follows that Universe Three would give the best future experience since it involves the least change in consumption levels and thus the lowest adjustment cost. If the respondents think that there are no consumption commitments—may be due to that they rent durable goods—and, therefore, that they pay no adjustment cost, they would choose the five universes uniformly. Both of these two patterns are not supported by the responses in which most respondents choose Universe One, and, therefore, adjustment cost cannot explain habit formation.

Cognition costs also cannot generate the response pattern. If one models cognition costs as associated with coming up with the optimal spending path, the respondents should choose the five universes uniformly because the consumption path is already specified in each universe and therefore require no cost to come up with it. If one models cognition costs as the cost associated with changing consumption path,²⁶ the respondents should choose Universe Three over others because it involves the least change of spending and, therefore, the least level of cognition costs. Whichever way one models cognition costs, the responses would be dramatically different from mine.

As a clarification of the above discussion, the fact that my evidence supports the existence of habit formation being separate from adjustment and cognition costs does not reject the existence of adjustment and cognition costs. What is shown is that habit formation exists separately from adjustment and cognition costs, and, consequently,

²⁵This is implied by the analyticity of the utility function (Debreu, 1972).

²⁶E.g., one may need to spend some cognition efforts to adjust her spending to a level different from her previous level.

Table 5: Estimates of Habit Depreciation Speed

	MAP	Mean	Median	95% HPDI
Habit Depreciation Rate	1.11	1.11	1.11	[0.89, 1.31]
Habit Depreciation Factor	0.67	0.67	0.67	[0.60, 0.74]

Note: The habit depreciation factor is annual.

that the latter cannot be deeper causes of the former.

4.2 Habit Depreciation Speed

To find out how fast does habit depreciate, I elicit and estimate the parameter θ as in $\dot{H} = \theta(C - H)$. In the survey question (Figure 3), I vary the time structure of past spending so that the resulting difference in current levels of habit enables the elicitation of θ (Proposition 1).

Proposition 1. *Under exact elicitation, $\theta > -\ln\left(1 - \frac{\Delta C_{U1}}{\Delta C_{U2}}\right)$ if the respondent chooses Universe One over Universe Two for a better future experience in the habit depreciation rate question.*

All proofs are in section D of the appendix.

In Proposition 1, ΔC_{U1} (ΔC_{U2}) is the difference between the monthly spending in Universe One (Universe Two) and the baseline monthly spending, \$5,000 per month. In the example of Figure 3, $\Delta C_{U1} = \$2,000$ and $\Delta C_{U2} = \$4,000$. Thus, according to Proposition 1, this survey question separates the values of θ into two complementary intervals: $\theta > \ln 2$ and $\theta < \ln 2$.²⁷ I use unfolding brackets to further pin down a finer range of θ for each possible response. Each respondent answers one to two follow-up questions that put her answer into one of the six brackets in Figure 7.

I estimate a depreciation rate of 1.11, which implies that habit depreciates 67% annually (Table 5). The annual depreciation rate of 67% implies that habit will have changed about 90% after two years, which is remarkably close to the findings in the psychology literature that income adaptation takes about two years.²⁸

The speed at which habit depreciates matters. In a typical additive habit formation model, the faster habit depreciates, the less risk-averse agents of the model become because habit adjusts faster with consumption. The most cited paper on habit formation, Campbell and Cochrane (1999),²⁹ also employs the additive habit.³⁰

²⁷I abstract from $\theta = \ln 2$ because θ has a probability of 0 to be exactly equal to $\ln 2$. The threshold of $\ln 2$ in continuous time corresponds to a threshold of 0.5 in discrete time at the annual frequency.

²⁸See Clark, Frijters and Shields (2008) for a review.

²⁹Google scholar reports that this paper has been cited by 4,925 times as of November 3, 2019.

³⁰They specified a nonlinear evolution for habit, surplus consumption ratio to be precise. It coincides with the linear habit evolution specified here in steady state.

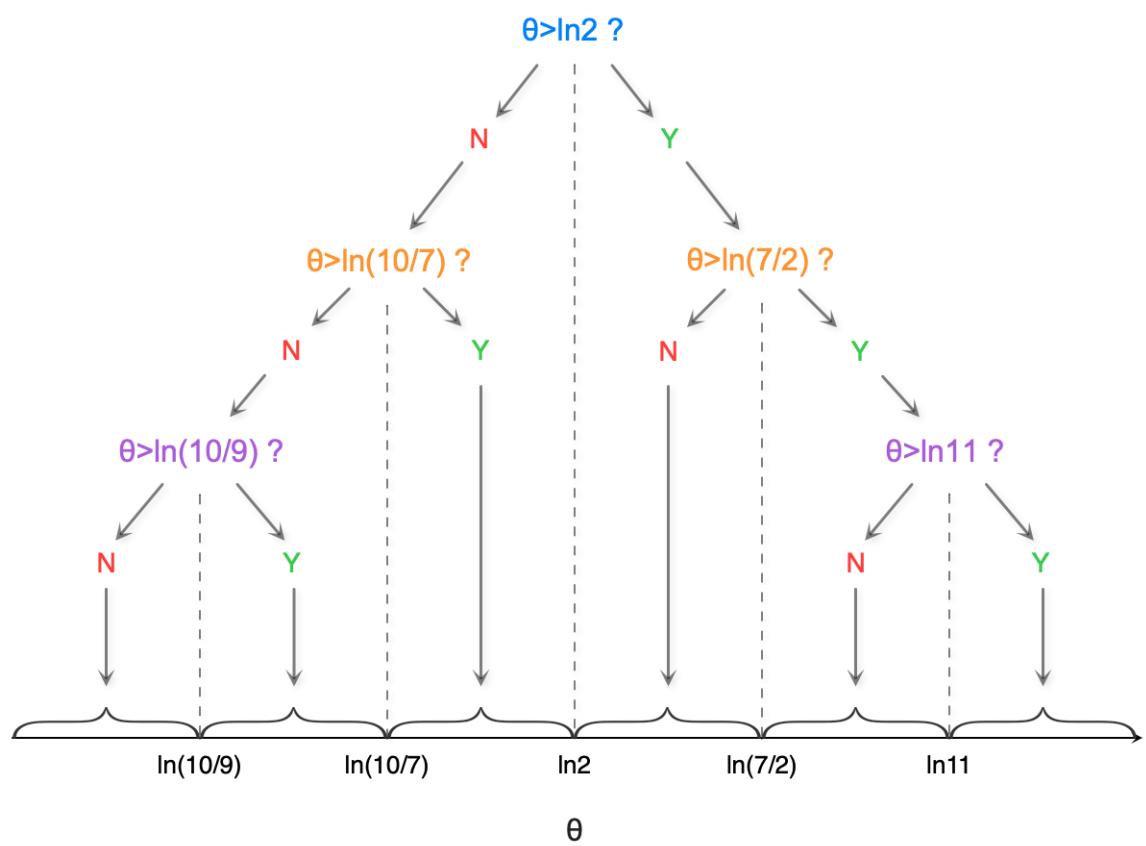


Figure 7: Unfolding Brackets

In their model, however, the faster habit depreciates, the more risk-averse people are. The reason is that in their model, the implied steady-state habit intensity³¹ is not constant, but increases with the habit depreciation rate. The higher the habit intensity is, the more likely that a fluctuation of consumption causes consumption to fall below the habit intensity adjusted level of habit, and thus the more risk-averse the agent becomes. The net effect of a higher habit depreciation rate in their model is the sum of these two effects, which turns out to be making the agent in the model more risk-averse.

Plugging my estimate of the habit depreciation speed into Campbell and Cochrane (1999) makes the agents become so risk-averse that they require such a high equity premium that has not been observed in history (Table 6). The time discount factor also has to be unrealistically low, 0.35 per year, to match the mean historical risk-free rate (column (3)). When a more realistic annual time discount factor of 0.89 is used, the level Campbell and Cochrane (1999) chooses, people become even more risk-averse. They require a still higher expected return and are willing to accept a hugely negative interest rate, -92.19% per year, to be able to save (column (4)). The intuition is that when the higher time discount factor makes people care more about the future, future risk matters more to them, and, as a result, they become yet more risk-averse. The higher risk aversion drives up the motive of precautionary saving. This motive is so strong that people are willing to pay more than 90% of the principal to be able to transfer the remaining less than 10% of it to the next year. When one lowers the time discount factor or the habit depreciation factor, the model moments are closer to reality, but the percentage differences are still at least 40% (columns (5) and (6)), even when habit depreciates only 30% each year which is very far from the 99% HPDI of the habit depreciation factor.

Table 7 shows the effect of my estimated depreciation factor on the predictability of dividend yield on expected return. Under my estimated habit depreciation factor, the regression coefficients and R^2 's are larger than post-war counterparts. It is also interesting to see that R^2 's start to decrease after the second year. Figure 8 replicates Figure 9 of Campbell and Cochrane (1999), with an additional curve showing model predicted P/D under the habit depreciation factor of 0.67. Even though the fluctuations retain about the same pattern, a higher habit depreciation speed makes P/D much smaller. This observation is consistent with that a higher value of θ induces agents to be more risk-averse and, therefore, to require a higher return to take on the same level of consumption risk.

My respondents might not be representative of marginal investors who price the assets. It is, however, unnecessary for my respondents to represent the marginal investors in every way possible. The above discussion stands as long as the typical habit depreciation speed of my respondents is the same or close to that of the typical

³¹Under additive habit, the instantaneous utility is $u(C - \alpha H)$ where α is the habit intensity parameter. In Campbell and Cochrane (1999)'s notation $X_t = \alpha_t H_t$. The implied steady-state habit intensity is, therefore, X_t/C_t .

Table 6: Effect of Habit Depreciation Speed in Campbell and Cochrane (1999): Equity Premium

	Postwar Sample		Model Sample			
	(1)	(2)	(3)	(4)	(5)	(6)
Habit Depreciation Factor	-	0.11	0.67	0.67	0.60	0.30
Time Discount Factor	-	0.89	0.35	0.89	0.42	0.71
Expected Excess log Return	6.69%	6.71%	44.08%	101.56%	37.44%	16.51%
Std of Excess log Return	15.20%	15.64%	31.08%	96.70%	29.62%	22.01%
Sharpe Ratio	0.43	0.43	1.42	1.05	1.26	0.75
Mean of Risk Free Rate	0.94%	0.94%	0.94%	-92.19%	0.94%	0.94%

Notes: All annualized values. Bold numbers are my changes to Campbell and Cochrane (1999)'s calibration. Column (1) is based on postwar (1947-95) value-weighted New York Stock Exchange stock index returns and 3-month Treasury bill rate; column (2) model sample based on Campbell and Cochrane (1999)'s calibration (0.11 is the annual habit depreciation factor implied by Campbell and Cochrane (1999)'s calibration of the persistence coefficient, ϕ , of the surplus consumption ratio in their model); column (3) model sample based on my estimate of habit depreciation factor; column (4) model sample based on my estimate of habit depreciation factor and the time discount factor of 0.89; column (5) model sample based on the lower bound of the 95% HPDI of my estimate of habit depreciation factor; column (6) model sample based on a habit depreciation factor far smaller than the lower bound of the 99% HPDI of my estimate of it. The power coefficient of the CRRA utility function is 2, as in Campbell and Cochrane (1999).

Table 7: Effect of Habit Depreciation Speed in Campbell and Cochrane (1999): Long-run Regression

Horizon (Years)	$\theta = 0.11$						$\theta = 0.67$							
	Postwar Sample			Consumption Claim		Dividend Claim		Consumption Claim			Dividend Claim			
	10 × Coefficient	R^2		10 × Coefficient	R^2		10 × Coefficient	R^2		10 × Coefficient	R^2			
1	-2.6	.18		-2.0	.13		-1.9	.08		-9.6	.60		-9.7	.53
2	-4.3	.27		-3.7	.23		-3.6	.14		-12.9	.71		-13.0	.60
3	-5.4	.37		-5.1	.32		-5.0	.19		-14.0	.67		-14.0	.56
5	-9.0	.55		-7.5	.46		-7.3	.26		-14.7	.55		-14.7	.44
7	-12.1	.65		-9.4	.55		-9.2	.30		-14.4	.44		-14.3	.34

Notes: Long-horizon regressions of log excess stock returns on the log price/dividend ratio in historical and simulated data. θ is annual habit depreciation factor. $\theta = 0.11$ is Campbell and Cochrane (1999)'s calibration and $\theta = 0.67$ my estimate. This table replicates part of Table 5 of Campbell and Cochrane (1999).

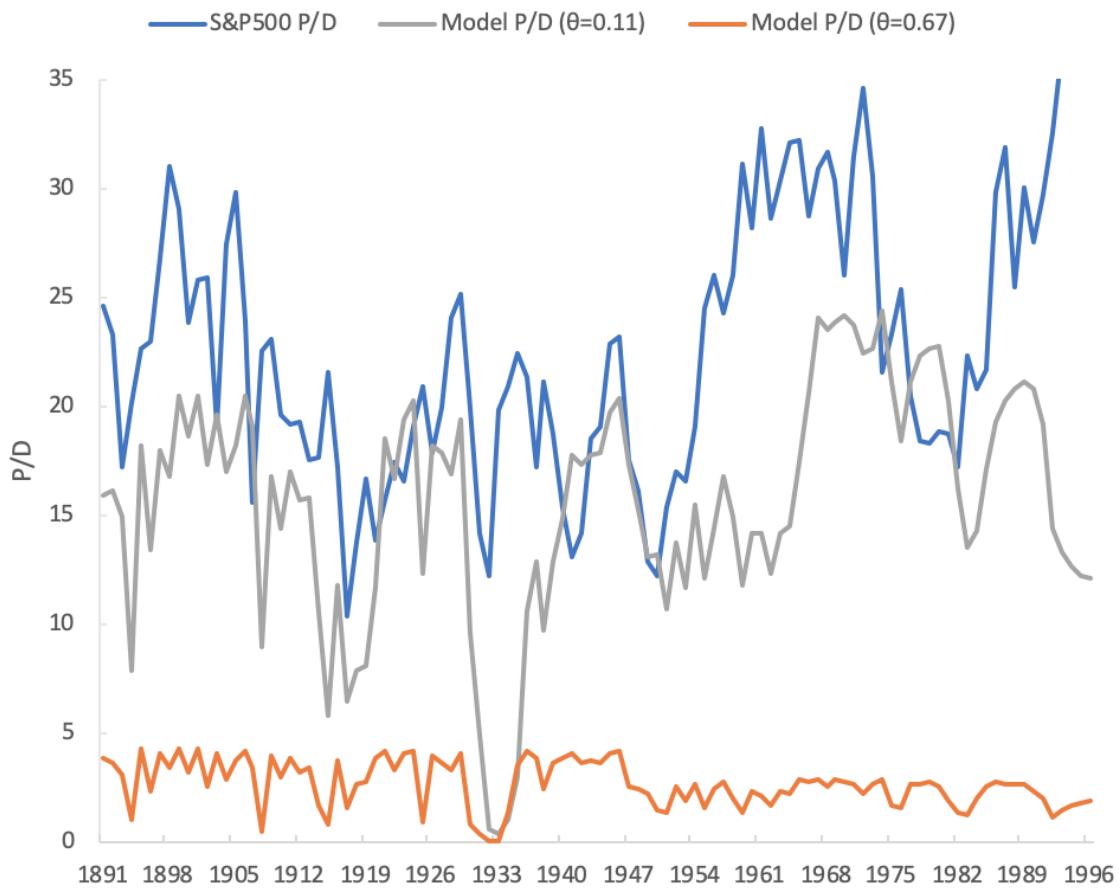


Figure 8: Historical Price/Dividend Ratio and Model Predictions Based on Consumption History

marginal investor, which would be the case if this parameter is a deep preference parameter that does not vary much across demographics. In section 6, I present such evidence: habit depreciation rate does not vary with age, gender, household size, and household income. One potential explanation could be that the speed that people's habit adjusts is determined genetically.

One of my goals for this study is to uncover the true capability of habit formation models. The literature has done a great job showing what the potential of the models can be. What is missing is that whether habit formation is *indeed* that powerful. I have found one inconsistency with reality in a popular habit formation model when I plug in my micro-based estimate of habit depreciation factor. One, however, needs to take extra caution in interpreting this inconsistency. Given that my evidence supports the existence of habit formation, it is more likely that the way that habit formation is modeled needs to be improved (more on this in the following sections) than that modeling habit formation is the wrong way to go. Just like we need features beyond diminishing marginal utility, no matter how realistic and fundamental it is, to be able to explain reality better, we might also need features *in addition to* habit formation to fully explain asset pricing and other phenomena.

4.3 Additive and Multiplicative Habits

Additive and multiplicative habits constitute the vast majority, if not all, of habit specifications that have been taken to data. Additive habit is adopted relatively more often in the literature for its time-varying risk aversion while multiplicative habit, at least in its simplest form, is not able to do so. Does micro evidence support this theoretical choice? I answer this question here, starting with four tests that distinguish the two formulations.

Proposition 2. *Additive habit, $u(C, H) \equiv v(C - \alpha H)$ with $\alpha \in \mathbb{R}^+$, implies $\frac{u_{CH}}{u_{HH}} \frac{u_H}{u_C} = 1$ and $\frac{u_{CH}}{u_{CC}} \frac{u_C}{u_H} = 1$.*

The intuition for this set of tests is that under additive habit, the indifference curves are parallel straight lines so that moving in any direction in the indifference map will not change the slope of the indifference curves. The two tests are the two bases spanning all such movements: increase H alone and increase C alone (Figure 9).

Proposition 3. *Multiplicative habit, $u(C, H) \equiv v(C/H^\alpha)$ with $\alpha \in \mathbb{R}^+$, implies $\frac{H u_H u_{CH}}{u_C u_H + H u_C u_{HH}} = 1$ and $\frac{C u_G u_{CH}}{u_C u_H + C u_H u_{CC}} = 1$.³²*

In the space of $(\ln C, \ln H)$, the two tests of Proposition 3 have similar intuition as those for the additive habit (Figure 10).

³²One can derive a class of tests based on the homotheticity of multiplicative habit. The two tests here imply this class of tests.

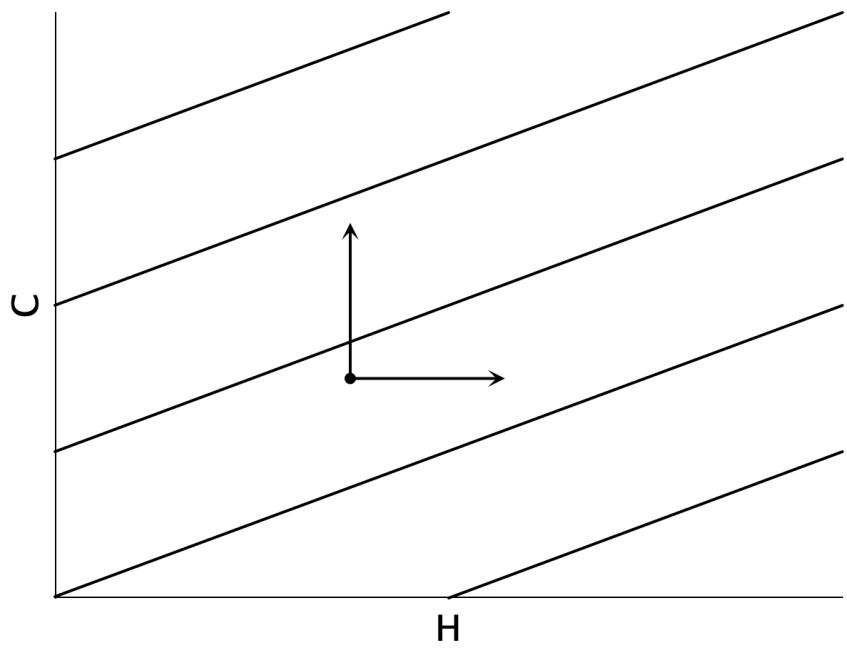


Figure 9: Additive Habit Indifference Map

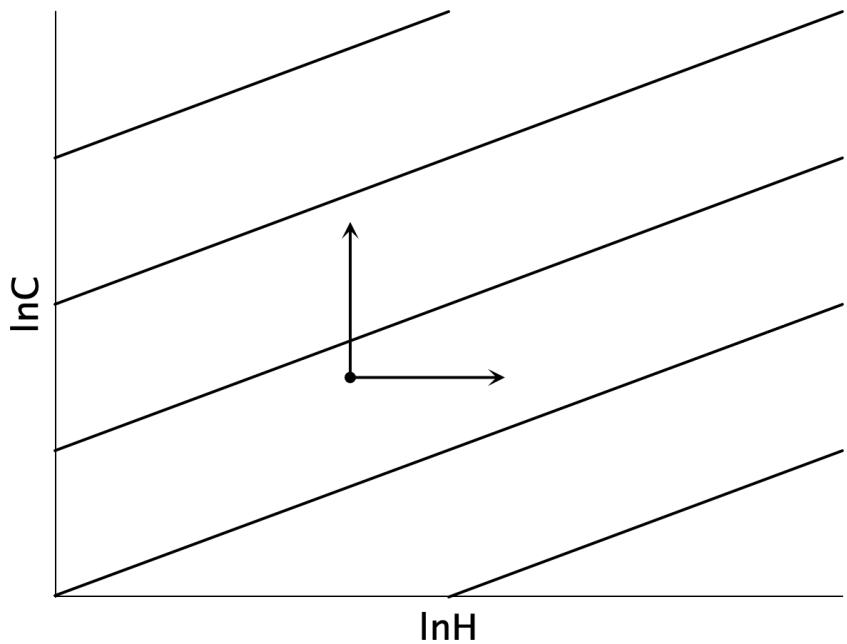


Figure 10: Multiplicative Habit Indifference Map in $(\ln C, \ln H)$ Space

Table 8: Estimates of Ratios of Utility Derivatives

	MAP	Mean	Median	95% HPDI
$-u_H/u_C$	0.62	0.62	0.62	[0.51, 0.73]
Hu_{HH}/u_H	7.71	7.78	7.76	[6.84, 8.78]
u_{CH}/u_{HH}	-0.86	-0.85	-0.86	[-1.00, -0.70]
u_{CC}/u_{HH}	3.72	3.72	3.72	[3.01, 4.41]

Because the left-hand sides of the four equations are functions of $-\frac{u_H}{u_C}$, $\frac{Hu_{HH}}{u_H}$, $\frac{u_{CH}}{u_{HH}}$, and $\frac{u_{CC}}{u_{HH}}$, implementation of the tests requires their elicitation. Due to the generality of the utility function, elicitation of these preference parameters will be up to third-order.

$-\frac{u_H}{u_C}$ is the slope (or inverse of the slope) of indifference curve. To elicit it, I change both future and past spending in the same direction to move along an indifference curve. The resulting survey question has monthly spending graphs as in Figure 11.

Proposition 4. *Under second-order elicitation, $-\frac{u_H}{u_C} < \frac{(\rho+\theta)\Delta f}{\rho\Delta e + \theta\Delta f}$ if the respondent chooses Universe One over Universe Two for a better future experience in the slope of indifference curve question.³³*

My estimate for the slope of indifference curve is 0.62 (Table 8). The implied positive sign of u_C is consistent with my assumption of the positive monotonicity of consumption. The magnitude of this estimate implies that about 62% of utility gain from a higher level of consumption gradually disappears because of habit formation. This number is remarkably close to the finding that about 60% of the effect of income on happiness is lost across time (Van Praag, 1971; Van Praag and Frijters, 1999).

Since the habit intensity parameter, α , equals the slope of indifference curve in both additive habit (everywhere) and multiplicative habit (in steady state), my estimate of 0.62 is essentially the same as the macro estimates of about 0.6 in the literature (Havranek, Rusnák and Sokolova, 2017). Havranek, Rusnák and Sokolova (2017) also finds that the estimates using microdata tend to be lower, 0.1, than those estimated using macrodata, 0.6. It is very interesting to see that my estimate using microdata is indistinguishable from those using macrodata.

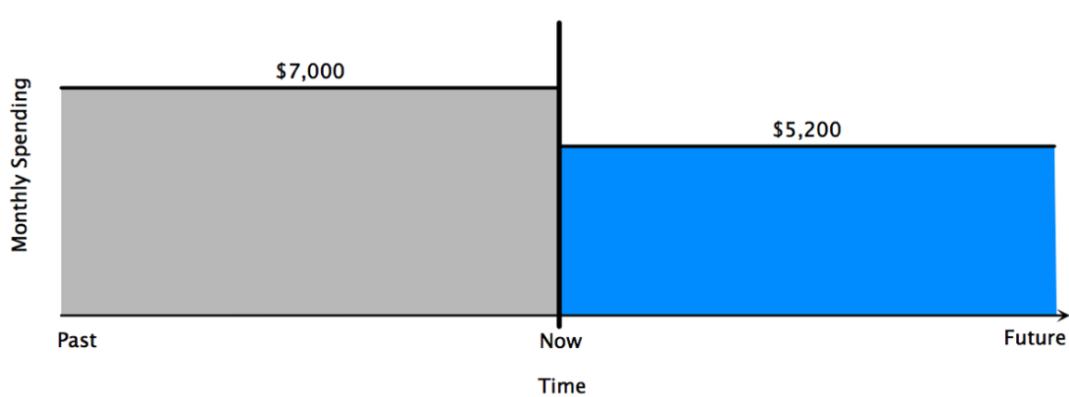
To elicit $\frac{Hu_{HH}}{u_H}$, I present the respondents with a tradeoff between past spending and fluctuation of past spending (Figure 12).

Proposition 5. *Under second-order elicitation, $\frac{Hu_{HH}}{u_H} < \frac{2(\rho+2\theta)}{\rho+\theta} \frac{\Delta f/\Delta e - 1}{(\Delta f/\Delta e)^2 + 1} \frac{H}{\Delta e}$ if the respondent chooses Universe One over Universe Two for a better future experience in the $\frac{Hu_{HH}}{u_H}$ question.³⁴*

³³ $\Delta e = \$2,000$ and $\Delta f = \$200$ in the example of Figure 11. Knowledge about the time discount rate, ρ , is required to estimate the slope of indifference curve and other preference parameters, the elicitation of which is relegated to section C of the appendix because of this indirect interest in it.

³⁴ $\Delta e = \$500$ and $\Delta f = \$650$ in the example of Figure 12.

Universe One



Universe Two

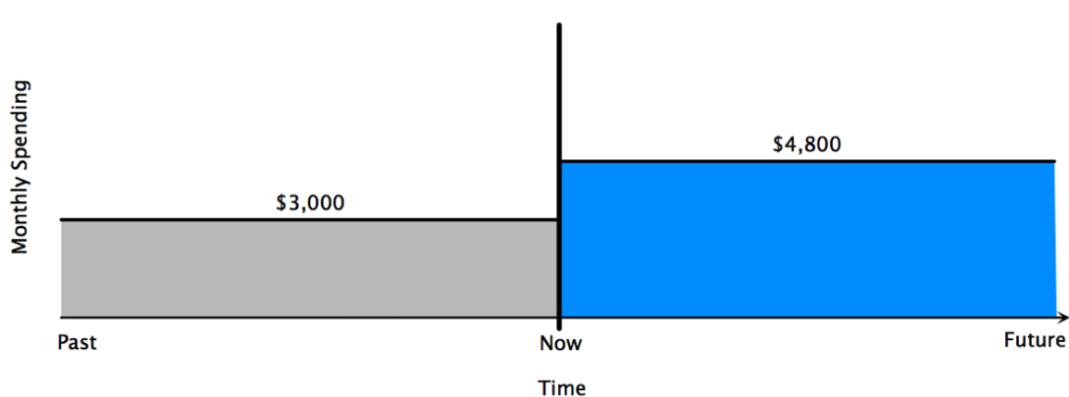


Figure 11: Monthly Spending Graphs - Slope of Indifference Curve

Universe One: 50-50 chance that your monthly spending is either Monthly Spending 1 or Monthly Spending 2 below



Universe Two: for sure that your monthly spending is

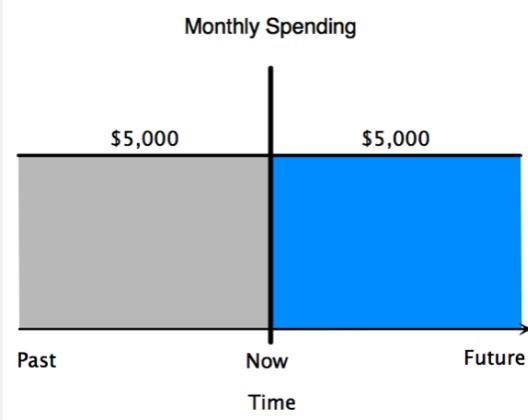


Figure 12: Monthly Spending Graphs - Hu_{HH}/u_H

Table 9: Estimates for Testing Additive and Multiplicative Habits

	MAP	Mean	Median	99% HPDI
$\frac{u_{CH}u_H}{u_{HH}u_C}$	0.53	0.53	0.53	[0.36, 0.71]
$\frac{u_{CH}u_C}{u_{CC}u_H}$	0.37	0.38	0.37	[0.23, 0.56]
$\frac{Hu_Hu_{CH}}{u_C(u_H+Hu_{HH})}$	0.46	0.47	0.47	[0.32, 0.63]
$\frac{Cu_Cu_{CH}}{u_H(u_C+Cu_{CC})}$	-0.24	-0.25	-0.24	[-0.34, -0.16]

I estimate $\frac{Hu_{HH}}{u_H}$ to be 7.71 (Table 8). Since I already estimated that $u_H < 0$, this implies that $u_{HH} < 0$.

The elicitation of $\frac{u_{CH}}{u_{HH}}$ rests on inducing fluctuations in both future and past spending at the same time. The monthly spending graphs for the survey question are in Figure 13.

Proposition 6. *Under third-order elicitation, $\frac{u_{CH}}{u_{HH}} < -\frac{(\rho+\theta)\Delta e + 2\theta\Delta f}{2(\rho+2\theta)\Delta f}$ if the respondent chooses Universe One over Universe Two for a better future experience in the $\frac{u_{CH}}{u_{HH}}$ question.³⁵*

My estimate of $\frac{u_{CH}}{u_{HH}}$ is -0.86 (Table 8). Given that I estimated that $u_{HH} < 0$, this tells me that $u_{CH} > 0$, which is consistent with the idea of habit formation: the higher habit is, ceteris paribus, the more valuable an additional unit of consumption is.

$\frac{u_{CC}}{u_{HH}}$ is about the tradeoff between two sources of fluctuations, one from future spending and the other from past spending, which leads to the monthly spending graphs in Figure 14.

Proposition 7. *Under third-order elicitation, $\frac{u_{CC}}{u_{HH}} < \frac{\rho}{\rho+2\theta} \left(\frac{\Delta e}{\Delta f} \right)^2 - \frac{2\theta}{\rho+\theta} \frac{u_{CH}}{u_{HH}} - \frac{2\theta^2}{(\rho+\theta)(\rho+2\theta)}$ if the respondent chooses Universe One over Universe Two for a better future experience in the $\frac{u_{CC}}{u_{HH}}$ question.³⁶*

$\frac{u_{CC}}{u_{HH}}$ is estimated to be 3.72 (Table 8). This is consistent with my assumption of $u_{CC} < 0$.

With the estimates of $-\frac{u_H}{u_C}$, $\frac{Hu_{HH}}{u_H}$, $\frac{u_{CH}}{u_{HH}}$, and $\frac{u_{CC}}{u_{HH}}$, I can calculate the statistics for testing additive and multiplicative habits. The point estimates show that none of the four statistics are equal to 1 (Table 9). Furthermore, 1 is not even in the 99% credible intervals of these statistics. These mean that my micro evidence supports neither additive nor multiplicative habit.

Given the existence of habit formation, there are two possibilities as to why this happened. On the one hand, the habit utility function is indeed of neither the additive

³⁵ $\Delta e = \$1000$ and $\Delta f = \$200$ in the example of Figure 13.

³⁶ $\Delta e = \$2200$ and $\Delta f = \$500$ in the example of Figure 14.

Universe One: 50-50 chance that your monthly spending is either Monthly Spending 1 or Monthly Spending 2 below



Universe Two: 50-50 chance that your monthly spending is either Monthly Spending 1 or Monthly Spending 2 below



Figure 13: Monthly Spending Graphs - u_{CH}/u_{HH}

Universe One: 50-50 chance that your monthly spending is either Monthly Spending 1 or Monthly Spending 2 below



Universe Two: 50-50 chance that your monthly spending is either Monthly Spending 1 or Monthly Spending 2 below

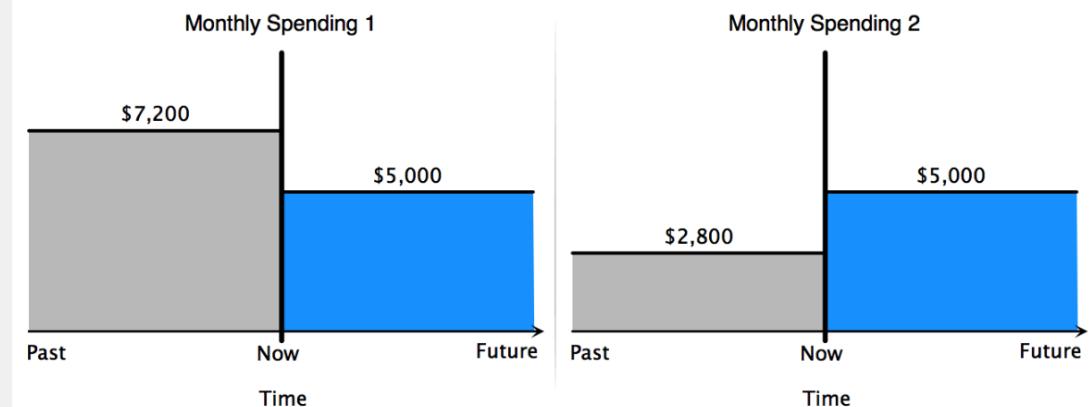


Figure 14: Monthly Spending Graphs - u_{CC}/u_{HH}

nor the multiplicative form. I have yet to see a third formulation of habit utility function that has been taken to data in the literature. If a micro-consistent form exists, it might be the key to solve the inconsistency I discussed in the last section and to explain phenomena current habit formation models have difficulty explaining. On the other hand, habit formation itself might be insufficient. The assumption of a representative agent might need to be relaxed or other features introduced to the model or both to be able to match the reality fully. The first possibility might be the near term direction for future research because my micro evidence is in line with the general idea of habit formation and does not exclude the possibility of habit preferences consistent with the evidence.

4.4 External Habit

So far, I have been abstracting from the effect of other people's past spending on my current well-being, the external habit. This type of habit formation has been prevalent in the literature, contrasting both the existing empirical evidence and our day to day intuition that the formation of habits originates mostly from my own behavior (internal habit). These two types of habits differ in their levels of tractability and can also have dramatically different implications for optimal tax policy and welfare analysis (Lettau and Uhlig, 2000; Ljungqvist and Uhlig, 2015). In this section, I explore the question of whether others' spending indeed affects my habit.

If an increase in others' past spending, *ceteris paribus*, makes me worse off in the future, external habit exists. That is, external habit exists if

$$Q_{EH} \equiv u(C, H(\{C\}, \{C_{others} + \Delta C_{others}\})) - u(C, H(\{C\}, \{C_{others}\})) < 0$$

for $\Delta C_{others} > 0$, where $\{\cdot\}$ denotes historical path, and I explicitly allow the dependence of habit on other people's past spending C_{others} .

To elicit the sign of Q_{EH} and the external habit mixture coefficient (see below), I use a set of questions where both others' and my past spending vary. This leads to the monthly spending graphs in Figure 15 and the elicitation in Proposition 8.

Proposition 8. *Given the existence of internal habit, $Q_{EH} < 0$ if the respondent chooses Universe One over Universe Two for a better future experience in the external habit question.*

The reverse of Proposition 8 does not hold: Choosing Universe Two over Universe One not necessarily implies $Q_{EH} \geq 0$. As long as external habit is weak enough relative to internal habit, a respondent would always prefer Universe Two for a better future experience. Under this scenario, my survey does not contain information that allows me to separate the different possibilities for the sign of Q_{EH} . Therefore, I check every possibility. If always choosing Universe Two means $Q_{EH} < 0$ (most lenient), all my response implies the existence of external habit. If always choosing Universe Two means $Q_{EH} > 0$ (strict), it is least likely to find evidence for external habit, while if

Universe One: your monthly spending graph is the left graph and others' monthly spending is the right graph below



Universe Two: your monthly spending graph is the left graph and others' monthly spending is the right graph below

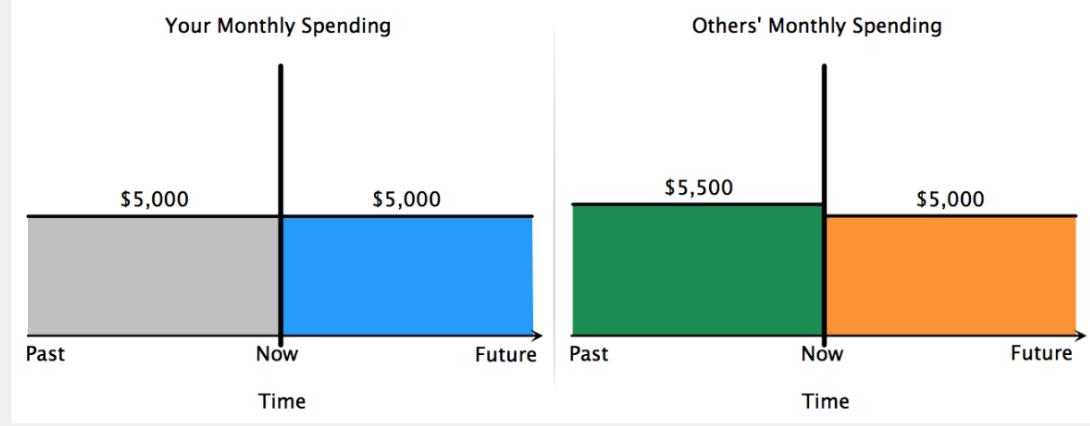


Figure 15: Monthly Spending Graphs - External Habit

Table 10: Response Frequency (Percentage) for External Habit Question

$\text{sgn}(Q_{EH})$	-1	-1, 0, or 1
First Wave	54	46
Second Wave	54	46

Table 11: Estimate of Existence of External Habit

	MAP	Mean	Median	90% HPMI	95% HPMI	99% HPMI
$\text{sgn}(Q_{EH})$ (strict)	-1.00	-0.84	-1.00	[-1.00, -1.00]	[-1.00, 1.00]	[-1.00, 1.00]
$\text{sgn}(Q_{EH})$ (lenient)	-1.00	-1.00	-1.00	[-1.00, -1.00]	[-1.00, -1.00]	[-1.00, -1.00]

Note: HPMIs of additional levels are shown to help determine confidence level.

always choosing Universe Two means $Q_{EH} = 0$ (lenient) it is relatively more likely to find external habit.

The response frequency in Table 10 indicates that the choice of the average respondent agrees with the existence of external habit. My estimate under the strict case shows that $Q_{EH} < 0$ in at least 90% HPMI while the opposite sign could emerge in 95% or higher-level HPMI (Table 11). The lenient scenario, however, finds evidence of a negative sign for Q_{EH} at the 99% level. Overall, external habit exists at least at the 90% level.

Composition of Internal and External Habits

Given the existence of internal and external habits, what is their relative importance? To provide an answer to this question, I augment the habit formation model. It is worth emphasizing that the augmentation of the model does not invalidate previous results because the augmented model nests the unaugmented model. Equivalently, one can think of my habit model as the most expanded version that includes everything that has been and will be discussed. I simply did not show some of the variables in previous sections because either they are constant or get canceled in elicitation, so not explicitly listing them results in no loss of generality.

I have already implicitly augmented the model in the first part of this section where I allow habit to depend on other people's spending. In this second part of the section, I specialize this dependence à la Grishchenko (2010):

$$\dot{H} = \theta((1 - \omega)C + \omega C_{others} - H) \quad (1)$$

where ω is the external habit mixture coefficient governing the contribution of others' spending (C_{others}) on my habit. To elicit this parameter, I can use the same question that elicits the existence of external habit in the first part of this section.

Proposition 9. *Under exact elicitation, $\omega > \frac{\Delta C}{\Delta C + \Delta C_{others}}$ if the respondent chooses*

Table 12: Estimate of External Habit Mixture Coefficient

	MAP	Mean	Median	95% HPDI
ω	0.17	0.16	0.16	[0.07, 0.26]

*Universe One over Universe Two for a better future experience in the external habit question.*³⁷

The point estimate indicates that others' spending contributes to about 17% of my habit (Table 12). The 95% HPDI of the external habit mixture coefficient does not include 0, which is consistent with the existence of external habit, as shown above.

4.5 Relative Strength of Habit Formation and Keeping Up with the Joneses

To elicit the relative strength of habit formation and keeping up with the Joneses, I unhide others' spending in the utility function, $u(C, H, C_{others})$. Then, I elicit the ratio of $\frac{u_{C_{others}}}{u_H}$ by varying others' spending in both the past and the future. The survey question has monthly spending graphs in Figure 16.

Proposition 10. *Under first-order elicitation, $\frac{u_{C_{others}}}{u_H} < \frac{\omega}{\rho+\theta} \left(\rho \frac{\Delta C_{others}^{U_2}}{\Delta C_{others}^{U_1}} - \theta \right)$ if the respondent chooses Universe One over Universe Two for a better future experience in the $\frac{u_{C_{others}}}{u_H}$ question.*³⁸

My point estimate for $\frac{u_{C_{others}}}{u_H}$ is about 1.05 (Table 13) and not significantly different from 1 at the 5% level. This supports that habit formation has about the same effect on utility as keeping up with the Joneses.

I draw two additional implications based on the significant negative sign of $u_{C_{others}}$ as implied from the estimate. First, keeping up with the Joneses exists separately from external habit. In the elicitation, external habit and keeping up with the Joneses are calculated separately. The fact that the estimate of $u_{C_{others}}$ is significantly negative means that keeping up with the Joneses exists after controlling for external habit. Second, keeping up with the Joneses is stronger than altruism. I did not restrict the sign of $u_{C_{others}}$ a priori, which can go both ways: altruism ($u_{C_{others}} > 0$) and keeping up with the Joneses ($u_{C_{others}} < 0$). Essentially $u_{C_{others}}$ represents the net effect of these two phenomena. The significant negative sign of $u_{C_{others}}$, therefore, indicates that keeping up with the Joneses dominates altruism.

³⁷ $\Delta C = \$500$ and $\Delta C_{others} = \$500$ in the example of Figure 15.

³⁸ $\Delta C_{others}^{U_1} = \300 and $\Delta C_{others}^{U_2} = \$3,000$ in the example of Figure 16.

Universe One: your monthly spending graph is the left graph and others' monthly spending is the right graph below



Universe Two: your monthly spending graph is the left graph and others' monthly spending is the right graph below



Figure 16: Monthly Spending Graphs - $u_{C_{others}}/u_H$

Table 13: Estimates for $u_{C_{others}}/u_H$

	MAP	Mean	Median	95% HPDI
$\frac{u_{C_{others}}}{u_H}$	1.05	1.06	1.06	[0.70, 1.41]

5 Explaining Easterlin Paradox

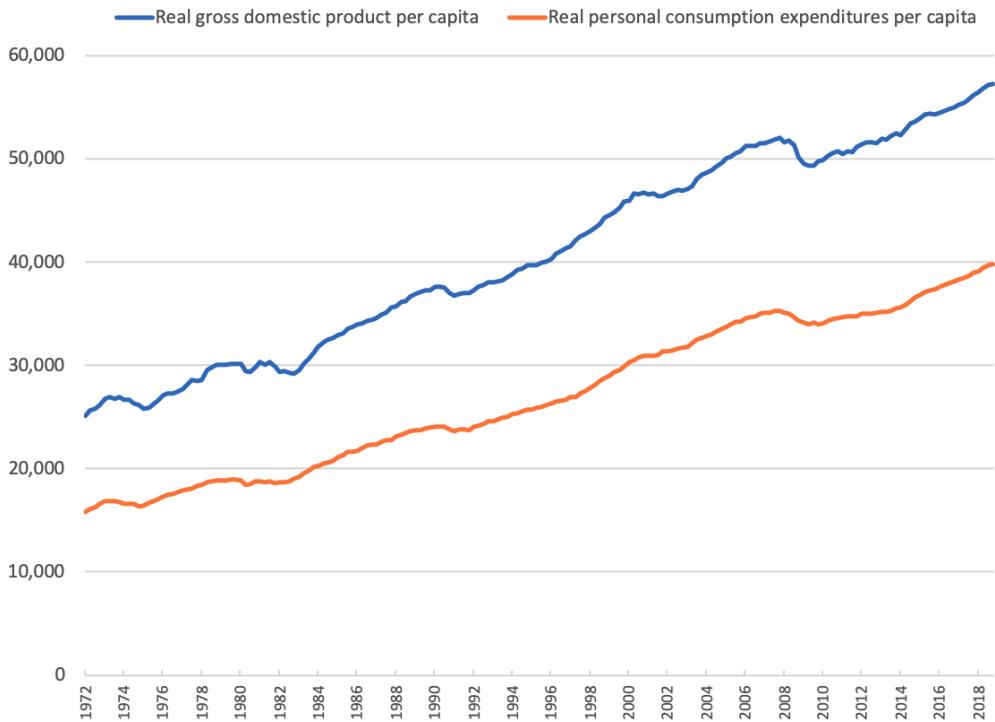
After nearly five decades of discussion, the income-happiness paradox proposed by Easterlin distills to the question that why income and happiness tend to be not associated in the long-run while positively associated in the short-run and cross-section (Easterlin, 1973, 1974, 1995; Stevenson and Wolfers, 2008; Sacks, Stevenson and Wolfers, 2012; Easterlin, 2017; Kaiser and Vendrik, 2018). Keeping up with the Joneses and habit formation have been proposed as potential explanations of this puzzle (Easterlin, 1973, 1974, 1995; Clark, Frijters and Shields, 2008). Recent evidence shows that keeping up with the Joneses alone cannot explain this phenomenon (Luttmer, 2005; Lewbel et al., 2018; De Giorgi, Frederiksen and Pistaferri, 2019). Using my estimates on keeping up with the Joneses and habit formation of both the internal and external types, I show in this section that while each alone cannot, joining forces, they can generate the patterns of income and happiness of the Easterlin Paradox.

Four clarifications merit discussion before proceeding. The first is that I focus on the causal channel that income changes happiness. Our typical life experiences and studies exploiting exogenous variations support income causing happiness to change (Frijters, Haisken-DeNew and Shields, 2004; Gardner and Oswald, 2007). Evidence aside, this causality motivated the discovery of the paradox³⁹ and is the most counter-intuitive, interesting,⁴⁰ and policy-relevant. Non-income happiness-altering factors do not help with the explanation of the paradox because they generally improve with income, making the long-run income-happiness relationship even more mysterious (Di Tella and MacCulloch, 2008). The second clarification is that following the literature (Clark, Frijters and Shields, 2008; Benjamin et al., 2012; Perez-Truglia, 2019), I assume the potential distinction between happiness and utility is of minimal effects on my discussion below. Third, the paradox holds when I replace income with consumption⁴¹: consumption moves closely with income (Figure 17), while happiness still has a long-term trend of about zero (Figure 18). The relative lack of attention to the relationship between consumption and happiness is at least partly due to a lack of reliable individual-level data on consumption. Compared to income, consumption relates more directly to human welfare, as is widely accepted in the economics literature. Fourth, there are at least three measures of happiness in the literature: Affect measures feelings of recent days, life satisfaction evaluation of life as a whole, and eudaemonia personal growth and meaning. I focus on the first two because their measurements are the most reliably measured (OECD, 2013), studied, and relevant

³⁹One can take a peek at the question that interested Easterlin from the titles of his seminal papers: “Does Money Buy Happiness?”(Easterlin, 1973), and “Does Economic Growth Improve the Human Lot? Some Empirical Evidence”(Easterlin, 1974).

⁴⁰This is evidenced by that the vast majority of speculated explanations of the paradox have focused on this channel.

⁴¹Throughout, consumption refers to real consumption.



Notes: Chained 2012 dollars. Data from U.S. Bureau of Economic Analysis.

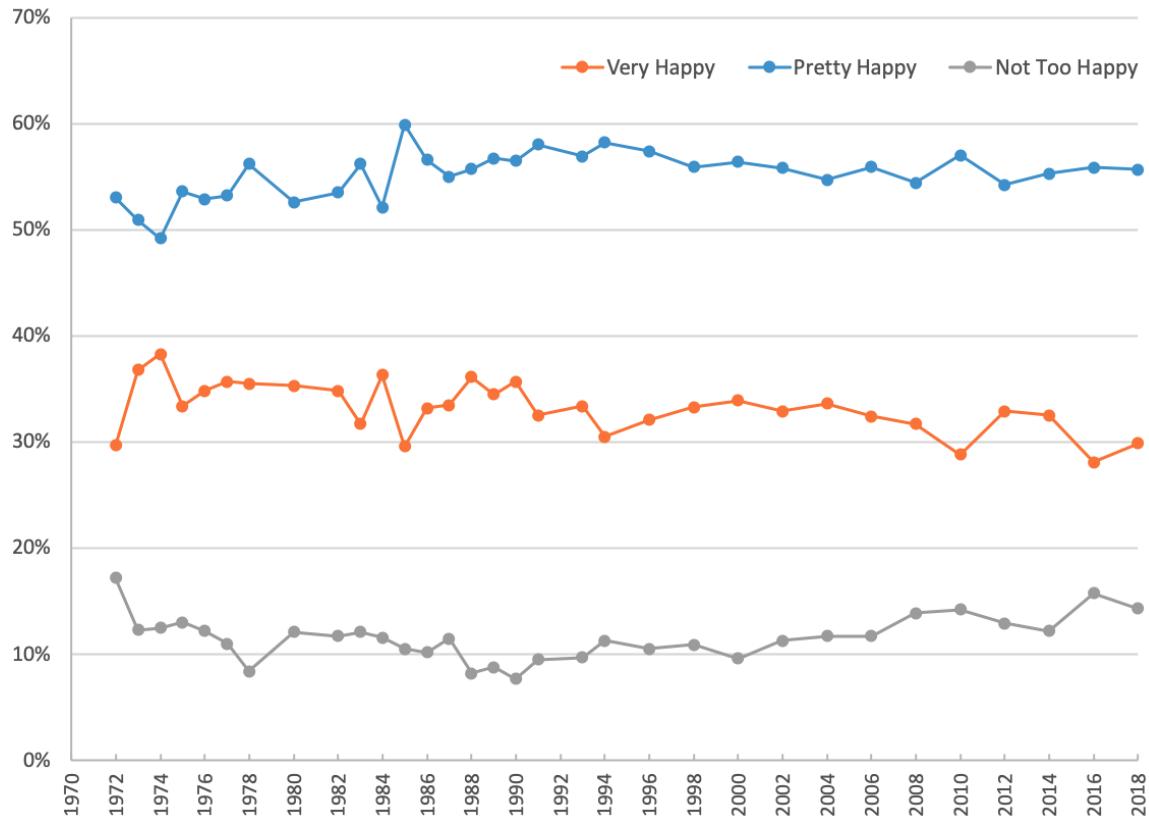
Figure 17: Real GDP and PCE Per Capita in the U.S., 1972 - 2018

for the paradox. I use the instantaneous utility as a proxy for affect⁴² and lifetime utility for life satisfaction.

To see how habit formation and keeping up with the Joneses potentially resolve the paradox, let me focus on a simplified environment that accentuates the mechanisms reconciling tensions of the paradox. I specify that people's utility is affected by both internal and external habits as well as keeping up with the Joneses. Habit evolves according to equation (1) with the habit depreciation rate and the external habit mixture coefficient calibrated to my estimates, 1.11 and 0.17, respectively. Keeping up with the Joneses and external habit emerge after the agent realizes that other people's consumption has changed, which is assumed to be k years after a universal consumption change happens.⁴³ With no loss of generality, I assume the full effect of keeping up with the Joneses applies instantly once others' spending changes become known to the agent.

⁴²One can alternatively use the integral of instantaneous utility over the past one day or week as proxies for affect, which are the typical time frames for survey questions on affect. I tried both and also other time frames and find that the difference is trivial. This is a necessary implication of the fact that the curvature of the instantaneous utility is rather smooth (see below).

⁴³The exact time of the delay does not matter for the resolution of the paradox. It only affects the speed that people's utility converges to its steady state.



Notes: Survey response from the survey question: "Taken all together, how would you say things are these days—would you say that you are very happy, pretty happy, or not too happy?" Data from General Social Surveys.

Figure 18: General Happiness in the U.S., 1972 - 2018

Table 14: Estimates for Easterlin Paradox

	MAP	Mean	Median	95% HPDI
$\frac{u_{C_{others}}}{u_C}$	-0.65	-0.65	-0.65	[-0.90, -0.41]
$\frac{u_H}{u_C} + \frac{u_{C_{others}}}{u_C}$	-1.26	-1.27	-1.27	[-1.60, -0.97]

The effects of habit formation and keeping up with the Joneses on utility, to a first-order approximation, is captured by u_H/u_C and $u_{C_{others}}/u_C$. My estimates of them are both less than -1 at the 95% level (Table 14), which means, contrary to common beliefs, each of them alone cannot fully explain the paradox.

The long-run zero income-happiness slope dictates that

$$\frac{u_H}{u_C} + \frac{u_{C_{others}}}{u_C} = -1 \quad (2)$$

, which my estimates support at the 95% level. The point estimate of the left-hand side is less than -1, which, aside from statistical precision considerations, provides the potential for my resolution to be consistent with the improvement of happiness-altering non-income factors (Di Tella and MacCulloch, 2008) and with the slightly negative long-run income-happiness slope in the U.S. (Stevenson and Wolfers, 2008; Firebaugh and Tach, 2012). Since my interest here is the link between income and happiness applicable not solely to the U.S., let me focus on the case where they sum to -1. For illustrative purposes, I choose $u_H/u_C = -0.55$ and $u_{C_{others}}/u_C = -0.45$, both of which are within their respective 95% HPDIs. As long as they sum to -1, the exact values of the two ratios only slightly affect the steady-state level of happiness and the convergence speed to steady states, both of which do not alter the income-happiness pattern that is what the Easterlin paradox is all about.

The intuition of equation (2) is that, to a first order approximation, habit formation and keeping up with the Joneses entirely cancel the happiness effect of permanent consumption changes in the long-run. To see this, suppose a single-episode-growth scenario where the economy was at some steady state such that its residents were at some constant level of happiness before the instant t_0 . Imagine that the economy grows at t_0 so that everyone's consumption permanently increases by a small amount of Δc for all instants starting from t_0 onwards (Figure 19a). As a result, to a first-order approximation,⁴⁴ the agent's affect goes up by $u_C \Delta c$ at t_0 . As time passes, people get used to this higher level of consumption, resulting in the buildup of internal habit that pulls affect down (Figure 19b). At the instant before keeping up with the Joneses kicks in, $t_0 + k$, the remaining effect of the higher consumption on affect is lowered to $[u_C + (1 - \omega)(1 - e^{-\theta k}) u_H] \Delta c$. At $t_0 + k$, the agent realizes that

⁴⁴Throughout my analysis in this section, I limit my attention to first-order approximations, as there is no point going beyond it when the value of $u_{C_{others}}/u_C$ is chosen based on its first-order elicitation. This is a good approximation when the change Δc is small, which I assume here.

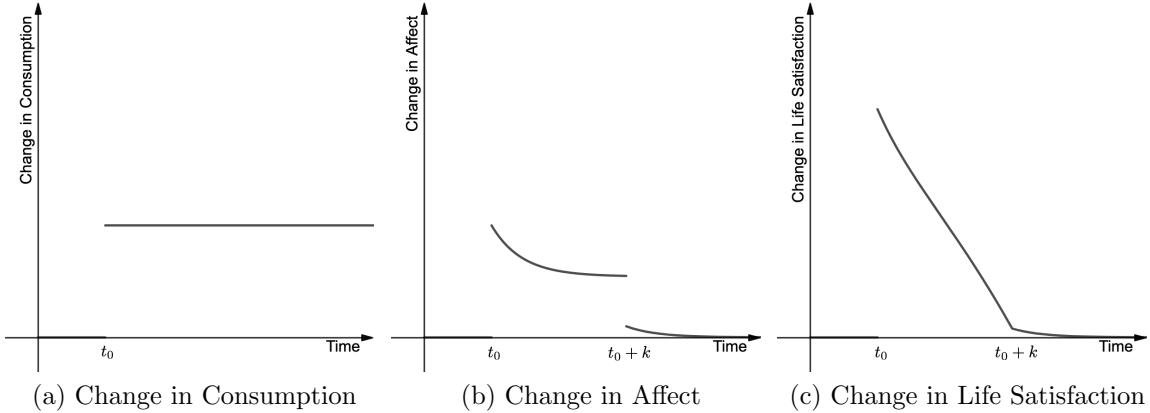


Figure 19: One-Episode Growth

everyone else also enjoys the same higher level of consumption as she enjoys and feels worse as a result of social comparison—keeping up with the Joneses—which further pushes the gain of affect down to $[u_C + (1 - \omega)(1 - e^{-\theta k}) u_H + u_{C_{others}}] \Delta c$. After that, external habit joins the play and, together with internal habit, gradually wears off the remaining gain of affect until it completely disappears.

Integrating the affect discounted by time preference,⁴⁵ I get the changes to life satisfaction (Figure 19c), the second measure of happiness. From the behavior of affect as analyzed above, it should come at no surprise that life satisfaction first increases then gradually decreases to its previous level. For later reference, let me call this pattern the wear-off effect: over time, habit formation and keeping up with the Joneses cancel out the happiness brought by a permanently higher level of consumption or income. Under the influence of this effect, at the distant future, the agent feels as if the permanent change to consumption did not happen even though it did. It is worth noting that before the wear-off effect reaches its full potency, higher consumption indeed brings higher levels of happiness. The wear-off effect reveals the key intuition of how habit formation and keeping up with the Joneses resolve the income-happiness paradox: the agent feels happier in the short-run after the consumption increase, but in the long-run this excitement fades away.

In reality, economies tend to grow over time, and, as a result, people typically earn more and consume more over time. To capture the key aspect of this phenomenon, suppose everyone increases her consumption permanently by Δc every one year after t_0 (Figure 20a). Figures 20b and 20c plot the changes to the agent’s happiness as time progresses. There should be of no surprise that habit formation wears off the gain of happiness within each year as in the one-episode-growth scenario above.⁴⁶ What

⁴⁵I calibrate the time discount rate to 0.15, based on my estimate of this parameter. See section C of the appendix for details.

⁴⁶The discontinuities of the utility changes at the start of each year after t_0 are purely out of the simplifying assumption that consumption permanently increases at the start of each year after t_0 ,

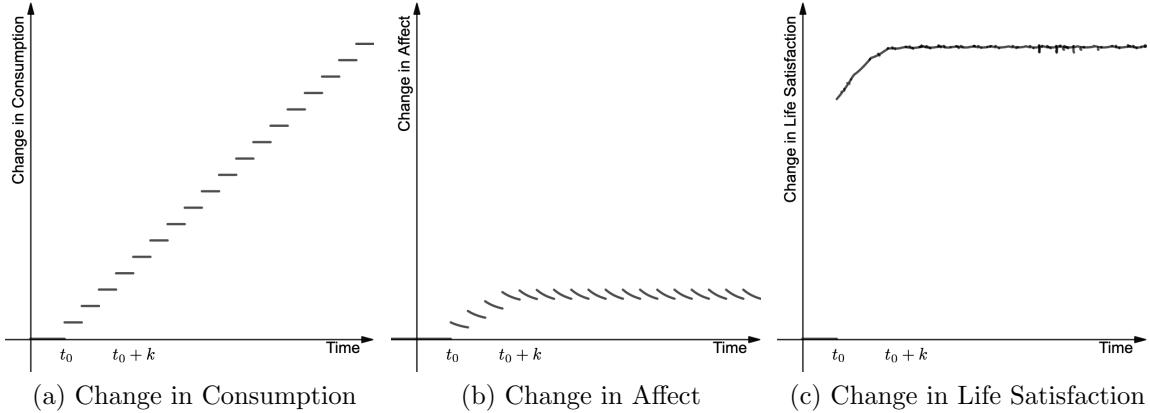


Figure 20: Multi-Episode Growth

is new is the dynamics of happiness: Instead of eventually going back to its previous level, happiness gradually builds up and then plateaus. Again, for later reference, let me label these two phases of the happiness dynamics transition effect and plateau effect. The transition effect exists, contrasting the decreasing trend as in the one-episode-growth scenario, because the growing consumption brings in each year a new episode of the wear-off effect whose initial happiness-enhancing phase⁴⁷ stacks onto those from previous years. As habit formation and keeping up with the Joneses take effect, the happiness-reducing momentum gradually builds up and eventually cancels out the happiness-enhancing momentum, leading us to the happiness plateau. The instant when such exact cancellation first happens is precisely the moment when the wear-off effect brought by the consumption increase at t_0 is in full swing.

Because the wear-off effect on happiness is proportional to Δc ,⁴⁸ the transition and plateau effects on happiness are also proportional to Δc (Figure 21). Let me call this the level effect: higher consumption growth leads to higher levels of happiness during both the transition and the plateau phases.

The level effect explains the cross-sectional positive correlation between income and happiness. Richer people or countries enjoy higher levels of consumption and therefore are on higher happiness curves, which implies that they are happier compared to poorer people or countries on the lower happiness curves.⁴⁹ Constant eco-

which is not essential. One can imagine that after an income increase, one increases consumption slowly, e.g., due to habit formation, so that the jumps in the utility changes disappear. All my analysis remain unchanged under this scenario. Here I do not employ this to highlight the wear-off effect of habit formation.

⁴⁷The time interval when happiness increases after t_0 in Figures 19b and 19c.

⁴⁸This is due to that my analysis here focuses on first-order approximations. To the extent that people's marginal utility of consumption is always positive, the analysis still holds: Even though the utility effect of the difference of high and low consumption changes will be smaller, the difference remains positive.

⁴⁹My above analysis has assumed a representative agent, purely to isolate the key mechanisms

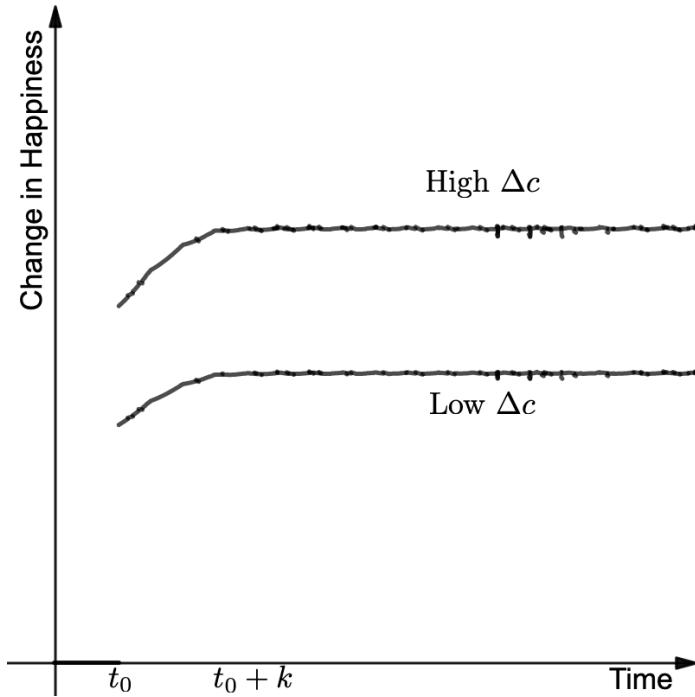


Figure 21: Level Effect

nomic fluctuations in the reality make consumption fluctuate, putting the agent frequently into transition phases. The transition effect, therefore, explains the short-run positive correlation between income and happiness. It is worth noting that regardless of income (or consumption) increase or decrease, transition effect always implies a positive relationship between income and happiness. Faster growing economies, such as developing countries, tend to enjoy faster economic growth, which by the level effect, leads to larger increases in happiness during the transition phase, as is observed in reality (Frijters, Haisken-DeNew and Shields, 2004). The plateau effect explains the long-run nil correlation between income and happiness. Even though income constantly fluctuates, it fluctuates around its trend. This trend growth determines the plateau level of happiness, which pins down the long-run trend of happiness. To put it another way, in the long-run, the trend of the happiness curve flattens even though consumption and income continually grow, hence the nil correlation.

To deepen the intuition, let me draw an analogy. Imagine you are about to run against a down escalator with a uniform speed— Δc stairs per unit of time. The

for resolving the paradox from potential complications implied in heterogeneities of the reference groups for social comparison. The analysis carries through with reasonable specifications of the reference groups. For example, one can assume that the income of the reference group changes by the same amount as the (heterogeneous) agent's income does. Note also that the level effect remains if everyone in the economy gets proportionally richer, implying that rising income inequality is not a requisite for resolving paradox with the help of habit formation as is suspected in the literature (Clark, 2016).

Table 15: Two Paradoxes

Time Frame	Easterlin Paradox	Escalator Paradox	Explanation
Long-run	Why getting more <i>income</i> doesn't raise happiness ?	Why running more <i>stairs</i> doesn't raise elevation ?	Plateau effect
Short-run	Why getting more <i>income</i> raises happiness ?	Why running more <i>stairs</i> raises elevation ?	Transition effect (+ fluctuation)
Cross-section	Why <i>richer</i> people/countries are happier ?	Why <i>faster</i> people are more elevated ?	Level effect

escalator initially is still and after you step onto it will gradually accelerate to the speed

$$\frac{u_H + u_{C_{others}}}{u_C} \Delta c = -\Delta c$$

stairs per unit of time. Suppose the escalator is long enough so that it will catch up to your speed before you can reach the other end of it. Your elevation represents happiness. The total stairs you run is the level of consumption, which implies that the speed you run, Δc stairs per unit of time, equals the consumption growth rate per unit of time, Δc per unit of time. The escalator concretizes the joint effect of habit formation and keeping up with the Joneses.

With this analogy, allow me to propose and resolve another paradox, the Escalator paradox, which parallels the Easterlin paradox (Table 15). The Escalator (Easterlin) paradox states that running more stairs (getting more income) raises elevation (happiness) in the cross-section and short-run but not so in the long-run. Why is this the case? In the long-run, the escalator (habit formation and keeping up with the Joneses) eventually catches up to your running speed (consumption growth), after which the additional stairs you run (additional consumption you enjoy) do not affect your elevation (happiness). In the short-run, you gain elevation (happiness) because your running speed (consumption growth) is faster than that of the escalator (the happiness canceling effect of habit formation and keeping up with the Joneses). In the cross-section, people who run faster (people or countries that are richer) are more elevated (happier) because the absolute difference between their speed (consumption growth) and the speed of the escalator (the happiness canceling effect of habit formation and keeping up with the Joneses) is larger.

How does the above discussion speak to the questions that motivated the paradox: does money buy happiness (Easterlin, 1973), and does economic growth improve human lot (Easterlin, 1974)? To phrase the questions in a slightly more accurate way, should we promote economic growth if happiness eventually will stop growing with economic growth? The answer is yes. Happiness would decrease if the economy

shrinks or grows at slower speeds. In other words, economic growth initially drives up happiness and eventually *maintains* it. If the economy grows slower or even shrinks, the resulting slower consumption growth will cause happiness to drop. Even though habit formation and keeping up with the Joneses will eventually bring it up a little bit, happiness will plateau at a level lower than the level it would plateau had the economy not slowed down.

6 Discussion and Robustness

In this section, I discuss the robustness of my results to departures from some assumptions underlying the above results.

6.1 Demographic Heterogeneity

I collected information on age, gender, household size, and household income of my respondents. It is interesting to see how my estimates vary with the demographics. Modifying the statistical model to allow the demographic variables to shift the means of the parameter distributions, I find that the demographics do not affect my estimates (Table 16). In particular, 0 is included in all of the 95% HPDIs of the estimated effects of demographic variables except that of household income on $\frac{Hu_{HH}}{u_H}$. After accounting for multiple hypothesis testing, this effect vanishes.⁵⁰

This result supports the view that the parameters I elicited are deep preference parameters that do not vary with sociodemographic characteristics. Since the ratios of utility derivatives are dependent on the spending profiles, the fact that the estimates do not vary with the demographics of the respondents and, therefore, their heterogeneous spending profiles in reality is reassuring, for it implies that my respondents understood the hypothetical situations of the survey and were able to answer the survey questions without letting their own demographic situations confound their responses.

6.2 Finite Horizon

I assume infinite horizon in my habit model, as do almost all current habit formation models in the literature. To investigate the effect of this assumption on my results, I re-derive all my elicitations of the preference parameters under finite horizon, resulting in minimal changes: no change for the elicitations of some parameters and tiny changes for the rest.⁵¹ I rerun my estimation under the finite-horizon elicitations and

⁵⁰The adjusted probability of the estimate greater than 0 under the Holm algorithm for this effect is 0.39.

⁵¹The thresholds for the habit depreciation rate and external habit mixture coefficient are exactly the same under both time horizons. The changes to the thresholds of other parameters are simply replacing 1 by $1 - e^{-\rho T}$ or $1 - e^{-(\rho+\theta)T}$ or $1 - e^{-(\rho+2\theta)T}$, all of which are close or very close to 1

Table 16: Effect of Demographics on Parameter Estimates

	Omitted Category	Age	Gender	Household Size	Household Income
Habit Depreciation Rate	1.31 [0.99, 1.65]	0.00 [-0.02, 0.01]	-0.24 [-0.67, 0.20]	0.13 [-0.03, 0.29]	-0.04 [-0.10, 0.01]
External Habit Mixture Coefficient	0.13 [0.00, 0.25]	0.00 [-0.01, 0.01]	0.06 [-0.13, 0.22]	-0.04 [-0.11, 0.03]	-0.02 [-0.04, 0.00]
$-u_H/u_C$	0.67 [0.50, 0.85]	0.00 [-0.01, 0.01]	-0.09 [-0.33, 0.14]	0.05 [-0.04, 0.13]	0.01 [-0.02, 0.04]
Hu_{HH}/u_H	6.62 [5.94, 8.12]	-0.02 [-0.09, 0.04]	1.00 [-0.02, 1.48]	-0.58 [-0.95, 0.02]	0.23 [0.03, 0.44]
u_{CH}/u_{HH}	-0.91 [-1.12, -0.68]	0.00 [-0.01, 0.01]	0.10 [-0.19, 0.40]	0.05 [-0.06, 0.16]	0.02 [-0.02, 0.06]
u_{CC}/u_{HH}	4.04 [2.98, 4.89]	0.03 [-0.03, 0.09]	-0.78 [-1.50, 0.46]	-0.32 [-0.86, 0.16]	-0.04 [-0.22, 0.12]
$u_{C^{others}}/u_H$	1.07 [0.52, 1.64]	-0.02 [-0.05, 0.01]	0.06 [-0.68, 0.81]	0.18 [-0.07, 0.46]	0.05 [-0.04, 0.14]
Time Discount Rate	0.09 [-0.09, 0.25]	0.00 [-0.01, 0.00]	0.12 [-0.09, 0.35]	0.02 [-0.06, 0.10]	0.00 [-0.02, 0.03]

Notes: The omitted category is that of 40-year-old males who live in 3-member households with \$50,001 - \$60,000 annual household income. 95% HPDI below MAP estimates.

Table 17: Finite Horizon Estimates

	MAP	Mean	Median	95% or 99%	HPDI/HPMI
Habit Depreciation Rate	1.12	1.11	1.11	[0.90, 1.32]	
External Habit Mixture Coefficient	0.16	0.16	0.16	[0.07, 0.26]	
$-u_H/u_C$	0.59	0.60	0.59	[0.49, 0.71]	
Hu_{HH}/u_H	7.66	7.77	7.75	[6.82, 8.73]	
u_{CH}/u_{HH}	-0.86	-0.85	-0.86	[-1.00, -0.71]	
u_{CC}/u_{HH}	3.83	3.85	3.85	[3.12, 4.57]	
$u_{C_{others}}/u_H$	1.15	1.14	1.14	[0.78, 1.51]	
$u_{C_{others}}/u_C$	-0.67	-0.68	-0.68	[-0.93, -0.42]	
$u_{CH}u_H/u_{HH}u_C$	0.49	0.51	0.51	[0.34, 0.69]	
$u_{CH}u_C/u_{CC}u_H$	0.37	0.38	0.37	[0.24, 0.56]	
$Hu_Hu_{CH}/u_C (u_H + Hu_{HH})$	0.45	0.45	0.45	[0.31, 0.61]	
$Cu_Cu_{CH}/u_H (u_C + Cu_{CC})$	-0.23	-0.24	-0.24	[-0.33, -0.16]	
Time Discount Rate	0.15	0.14	0.14	[0.04, 0.25]	

Note: The time horizon is 30 years in the future, as I instruct the respondents in the survey "... think of ... 'Future' as the next 30 years."

find that the estimates are almost the same as those in the infinite horizon (Table 17).

6.3 Additional Attention Checks

I have already utilized the explicit attention checks to screen out respondents who did not understand the hypothetical situation or the monthly spending graphs through quizzes and deleted the responses of those who sped the survey and were located outside of the U.S. In this section, I make use of implicit attention checks to see if a potential lack of attention biases my results. Because the implicit attention checks are not perfect proxies for attention, I apply them successively, from the relatively more reliable to the relatively less reliable.

At the end of the survey, I asked the quiz on the basic hypothetical situation of the survey again. There are 120 respondents in wave one and 54 in wave two who made at least one mistake in answering the five-question quiz. Deleting these responses from my sample leads to no significant change to my estimates (Table 18), except that external habit no longer exists under the strict case, though still exists under the lenient case.

I asked demographic questions in both waves. Within the relatively moderate amount of time that separates the two waves, the demographics should not change. In other words, the time consistency of answers to the demographic questions can serve

under reasonable values of T , the finite time horizon of interest.

Table 18: Estimates Using Subsample of No Failed Quiz

	MAP	Mean	Median	95% or 99% HPDI/HPMI
Habit Depreciation Rate	1.11	1.12	1.12	[0.86, 1.37]
External Habit Mixture Coefficient	0.08	0.10	0.10	[0.00, 0.21]
$-u_H/u_C$	0.69	0.70	0.70	[0.56, 0.84]
Hu_{HH}/u_H	8.21	8.30	8.27	[7.02, 9.61]
u_{CH}/u_{HH}	-0.80	-0.81	-0.81	[-1.00, -0.60]
u_{CC}/u_{HH}	4.45	4.42	4.42	[2.82, 5.97]
$u_{C_{others}}/u_H$	1.53	1.52	1.52	[0.80, 2.27]
$u_{C_{others}}/u_C$	-1.02	-1.07	-1.05	[-1.64, -0.52]
$u_{CH}u_H/u_{HH}u_C$	0.56	0.57	0.56	[0.33, 0.82]
$u_{CH}u_C/u_{CC}u_H$	0.24	0.27	0.26	[0.13, 0.51]
$Hu_Hu_{CH}/u_C (u_H + Hu_{HH})$	0.50	0.51	0.50	[0.30, 0.73]
$Cu_Cu_{CH}/u_H (u_C + Cu_{CC})$	-0.18	-0.20	-0.19	[-0.37, -0.10]
$\text{sgn}(Q_H)$	-1.00	-1.00	-1.00	[-1.00, -1.00]
$\text{sgn}(Q_{EH})$ (strict)	1.00	0.11	1.00	0.55
$\text{sgn}(Q_{EH})$ (lenient)	-1.00	-1.00	-1.00	[-1.00, -1.00]
Time Discount Rate	0.11	0.11	0.11	[-0.06, 0.26]

Notes: 99% HPDIs or HPMIs are reported for the seven statistics for testing additive and multiplicative habits and existence of habit formation. 95% HPDIs are reported for the rest parameters. In cases where an estimate is not significant at the 90% level, the probability of the estimate greater than 0 is shown.

Table 19: Estimates Using Subsample of No Demographic Mistake

	MAP	Mean	Median	95% or 99% HPDI/HPMI
Habit Depreciation Rate	1.15	1.13	1.13	[0.85, 1.41]
External Habit Mixture Coefficient	0.10	0.12	0.12	[0.00, 0.24]
$-u_H/u_C$	0.67	0.67	0.67	[0.52, 0.82]
Hu_{HH}/u_H	8.35	8.47	8.45	[7.27, 9.92]
u_{CH}/u_{HH}	-0.87	-0.86	-0.86	[-1.06, -0.65]
u_{CC}/u_{HH}	4.20	4.15	4.15	[2.47, 5.86]
$u_{C_{others}}/u_H$	1.45	1.44	1.43	[0.64, 2.22]
$u_{C_{others}}/u_C$	-0.91	-0.96	-0.94	[-1.54, -0.41]
$u_{CH}u_H/u_{HH}u_C$	0.57	0.57	0.57	[0.34, 0.83]
$u_{CH}u_C/u_{CC}u_H$	0.29	0.33	0.31	[0.14, 0.67]
$Hu_Hu_{CH}/u_C (u_H + Hu_{HH})$	0.51	0.51	0.51	[0.30, 0.75]
$Cu_Cu_{CH}/u_H (u_C + Cu_{CC})$	-0.20	-0.23	-0.22	[-0.46, -0.10]
$\text{sgn}(Q_H)$	-1.00	-1.00	-1.00	[-1.00, -1.00]
$\text{sgn}(Q_{EH})$ (strict)	-1.00	-0.31	-1.00	0.34
$\text{sgn}(Q_{EH})$ (lenient)	-1.00	-1.00	-1.00	99: [-1.00, -1.00]
Time Discount Rate	0.12	0.11	0.12	[-0.05, 0.27]

Notes: 99% HPDIs or HPMIs are reported for the seven statistics for testing additive and multiplicative habits and existence of habit formation. 95% HPDIs are reported for the rest parameters. In cases where an estimate is not significant at the 90% level, the probability of the estimate greater than 0 is shown.

as an implicit attention check. Applying this check deletes another 18 responses from the remaining sample. My estimates are essentially unchanged (Table 19) except, again, that the evidence for existence of external habit is weaker.

A third implicit attention check is that people should be indifferent toward the universes when there is no difference between them. In the time discount rate question⁵², past spendings are the same in both universes, where the respondents should choose the same past experience. This leads me to further delete 104 and 19 responses from waves one and two, respectively. The estimates remain in line with my baseline estimates (Table 20). Two points worth noting. The external habit exists at the 90% level in the strict case, as is in my main results. The credible intervals inflate as a result of the much smaller sample size. This reduces the precision of some of my statistics for the tests on the existence of habit formation and on additive and multiplicative habits. However, the rest of the statistics remain accurate enough to keep my results robust.

Finally, I use a measure of response consistency across the waves as an attention check. Considering this attention check involves more speculation, I delete only those

⁵²See section C of the appendix.

Table 20: Estimates Using Subsample of Same Past Experience

	MAP	Mean	Median	95% or 99% HPDI/HPMI
Habit Depreciation Rate	1.27	1.27	1.27	[0.84, 1.74]
External Habit Mixture Coefficient	0.27	0.27	0.26	[0.00, 0.50]
$-u_H/u_C$	0.49	0.50	0.50	[0.30, 0.69]
Hu_{HH}/u_H	9.12	8.78	8.87	[7.35, 10.00]
u_{CH}/u_{HH}	-0.69	-0.68	-0.68	[-0.95, -0.41]
u_{CC}/u_{HH}	4.33	4.41	4.39	[2.31, 6.57]
$u_{C_{others}}/u_H$	0.93	0.98	0.98	[-0.20, 2.12]
$u_{C_{others}}/u_C$	-0.44	-0.49	-0.47	[-1.11, 0.12]
$u_{CH}u_H/u_{HH}u_C$	0.32	0.34	0.33	[0.11, 0.60]
$u_{CH}u_C/u_{CC}u_H$	0.22	0.35	0.31	[0.06, 0.92]
$Hu_Hu_{CH}/u_C (u_H + Hu_{HH})$	0.28	0.30	0.30	[0.10, 0.54]
$Cu_Cu_{CH}/u_H (u_C + Cu_{CC})$	-0.14	-0.17	-0.16	[-0.44, -0.04]
$\text{sgn}(Q_H)$	-1.00	-0.97	-1.00	95: [-1.00, -1.00] 99: [-1.00, 1.00]
$\text{sgn}(Q_{EH})$ (strict)	-1.00	-0.85	-1.00	90: [-1.00, -1.00] 95: [-1.00, 1.00]
$\text{sgn}(Q_{EH})$ (lenient)	-1.00	-1.00	-1.00	99: [-1.00, -1.00]
Time Discount Rate	0.10	0.10	0.10	[-0.27, 0.46]

Notes: 99% HPDIs or HPMIs are reported for the seven statistics for testing additive and multiplicative habits and existence of habit formation. 95% HPDIs are reported for the rest parameters. Additional HPDI or HPMI of relevant level is also shown to help determine the confidence level.

Table 21: Estimates Using Sample of No Polar Response

	MAP	Mean	Median	95% or 99%	HPDI/HPMI
Habit Depreciation Rate	1.43	1.43	1.43	[0.95, 1.96]	
External Habit Mixture Coefficient	0.15	0.22	0.20	[0.00, 0.48]	
$-u_H/u_C$	0.54	0.53	0.53	[0.25, 0.78]	
Hu_{HH}/u_H	8.82	8.54	8.61	[6.94, 10.00]	
u_{CH}/u_{HH}	-0.73	-0.72	-0.73	[-0.99, -0.44]	
u_{CC}/u_{HH}	4.69	4.56	4.57	[1.91, 7.23]	
$u_{C_{others}}/u_H$	1.02	1.07	1.05	[-0.25, 2.44]	
$u_{C_{others}}/u_C$	-0.45	-0.56	-0.53	[-1.36, 0.20]	
$u_{CH}u_H/u_{HH}u_C$	0.36	0.38	0.37	[0.08, 0.71]	
$u_{CH}u_C/u_{CC}u_H$	0.10	0.34	0.31	[0.07, 1.41]	
$Hu_Hu_{CH}/u_C (u_H + Hu_{HH})$	0.33	0.34	0.33	[0.08, 0.65]	
$Cu_Cu_{CH}/u_H (u_C + Cu_{CC})$	-0.16	-0.19	-0.17	[-0.67, -0.03]	
$\text{sgn}(Q_H)$	-1.00	-0.99	-1.00	99: [-1.00, -1.00]	
$\text{sgn}(Q_{EH})$ (strict)	-1.00	-0.32	-1.00	0.34	
$\text{sgn}(Q_{EH})$ (lenient)	-1.00	-1.00	-1.00	99: [-1.00, -1.00]	
Time Discount Rate	0.02	-0.05	-0.02	[-0.62, 0.47]	

Notes: 99% HPDIs or HPMIs are reported for the seven statistics for testing additive and multiplicative habits and existence of habit formation. 95% HPDIs are reported for the rest parameters. Additional HPDI of relevant level is also shown to help determine the credible level. In cases where an estimate is not significant at the 90% level, the probability of the estimate greater than 0 is shown.

who give at least one polar response—any response that corresponds to the first (last) extreme range of parameter values in wave one and the last (first) in wave two. This check leads to the deletion of another 31 responses from both waves, resulting in further expansion of the credible intervals of my estimates, but, again, the estimates are not significantly different, and my results remain robust (Table 21).

6.4 Response Error with Non-Zero and Wave-Varying Mean

In the statistical model, I assume that the response error is zero-mean across both waves of the survey. Relaxing this assumption, I arrive at a statistical model with response errors of non-zero means that potentially vary across waves. Without loss of generality,⁵³ the joint distribution of parameter \tilde{x} for individual i in both waves

⁵³Only two means, one for each wave, can be identified. The specification here identifying the difference of the means is equivalent to a specification that specifies the two means using two parameters, one for each mean. If the two means are different, μ_ϵ should be statistically different from 0.

Table 22: Means of Response Errors in Wave One

μ_ε of	MAP	Mean	Median	95% HPDI
Habit Depreciation Rate	-0.01	-0.02	-0.02	[-0.25, 0.21]
External Habit Mixture Coefficient	-0.05	-0.04	-0.04	[-0.13, 0.06]
$-u_H/u_C$	-0.01	-0.01	-0.01	[-0.10, 0.07]
Hu_{HH}/u_H	-0.85	-1.15	-1.03	[-3.06, 0.48]
u_{CH}/u_{HH}	0.08	0.08	0.08	[-0.07, 0.23]
u_{CC}/u_{HH}	-0.41	-0.41	-0.40	[-1.14, 0.34]
$u_{C^{others}}/u_H$	-0.13	-0.13	-0.13	[-0.52, 0.25]
Time Discount Rate	-0.05	-0.05	-0.05	[-0.16, 0.07]

Note: μ_ε denotes mean of response error.

becomes

$$\begin{bmatrix} \tilde{x}_{i,1} \\ \tilde{x}_{i,2} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_x + \mu_\varepsilon \\ \mu_x - \mu_\varepsilon \end{bmatrix}, \begin{bmatrix} \sigma_x^2 + \sigma_{\varepsilon_x}^2 + \sigma_{\varepsilon_{x,1}}^2 & \sigma_x^2 + \sigma_{\varepsilon_x}^2 \\ \sigma_x^2 + \sigma_{\varepsilon_x}^2 & \sigma_x^2 + \sigma_{\varepsilon_x}^2 + \sigma_{\varepsilon_{x,2}}^2 \end{bmatrix} \right)$$

The estimates of the means of the response errors in wave one or equivalently negative of the means of the response errors in wave two are indistinguishable from zero at the 5% level (Table 22). This implies that my estimates are robust to the non-zero-wave-varying-mean response errors, which is exactly what I find (Table 23).

7 Conclusion

In this paper, I provide an extensive set of survey evidence for habit formation. I find that people's spending behavior exhibits habit formation, partly explaining why adding habit formation tends to improve the explanatory power of models. Considering that there are phenomena current habit formation models are having a hard time matching, I suggest that future models should be built onto habit formation rather than ignoring it, especially when it alone does not seem to be able to match data.

Both internal and external habits exist, with the latter accounting for a small fraction (17%) of habit. This implies that in terms of micro evidence, internal habit is a better choice than external habit. Better still, a composition of internal and external habits with the estimated weights could potentially deliver an even superior match with people's behavior.

Habit depreciates about two-thirds per year. Future habit formation models should match or at least check how the model performs under this speed of habit depreciation. Specifying this parameter to the value can potentially reduce the fit

Table 23: Estimates under Non-Zero and Wave-Varying Means of Response Errors

	MAP	Mean	Median	95% or 99% HPDI/HPMI
Habit Depreciation Rate	1.11	1.12	1.12	[0.86, 1.36]
External Habit Mixture Coefficient	0.20	0.19	0.19	[0.07, 0.31]
$-u_H/u_C$	0.63	0.63	0.62	[0.51, 0.75]
Hu_{HH}/u_H	8.55	8.81	8.69	[7.03, 10.85]
u_{CH}/u_{HH}	-0.90	-0.89	-0.89	[-1.05, -0.72]
u_{CC}/u_{HH}	4.03	4.05	4.05	[3.15, 5.00]
$u_{C_{others}}/u_H$	1.14	1.14	1.14	[0.74, 1.59]
$u_{C_{others}}/u_C$	-0.70	-0.71	-0.71	[-1.02, -0.42]
$u_{CH}u_H/u_{HH}u_C$	0.55	0.56	0.55	[0.37, 0.76]
$u_{CH}u_C/u_{CC}u_H$	0.34	0.36	0.35	[0.22, 0.56]
$Hu_Hu_{CH}/u_C (u_H + Hu_{HH})$	0.50	0.50	0.50	[0.33, 0.68]
$Cu_Cu_{CH}/u_H (u_C + Cu_{CC})$	-0.23	-0.23	-0.23	[-0.35, -0.15]
$\text{sgn}(Q_H)$	-1.00	-0.97	-1.00	95: [-1.00, -1.00] 99: [-1.00, 1.00]
$\text{sgn}(Q_{EH})$ (strict)	-1.00	-0.85	-1.00	90: [-1.00, -1.00] 95: [-1.00, 1.00]
$\text{sgn}(Q_{EH})$ (lenient)	-1.00	-1.00	-1.00	99: [-1.00, -1.00]
Time Discount Rate	0.18	0.18	0.18	[0.04, 0.34]

Notes: 99% HPDIs or HPMIs are reported for the seven statistics for testing additive and multiplicative habits and existence of habit formation. 95% HPDIs are reported for the rest parameters. Additional HPDI or HPMI of relevant level is also shown to help determine the credible level.

between model and data. I believe that signals that habit formation alone is not sufficient to fully explain the data at hand, necessitating the introduction of additional features to the model.

None of the current habit formation models is consistent with people's behavior because their specific utility functions fail to pass the four tests on their validities. This justifies a search for a habit utility function that matches the micro evidence, which potentially can explain phenomena current habit models cannot explain.

In terms of welfare effects, habit formation is as important as keeping up with the Joneses. Both external habit and keeping up with the Joneses exist in people's behavior. Keeping up with the Joneses dominates altruism.

Combining habit formation with keeping up with the Joneses can generate the patterns of income (consumption) and happiness highlighted by the Easterlin paradox. This implies that happiness can increase with income, but only for a while before the wear-off effect induced by habituation and social comparison ends the happiness-enhancing phase. Level and transition effects lead to the cross-sectional and short-run positive correlations of income and happiness while plateau effect the long-run nil correlation. Even though happiness only increases for a while, continued income growth is necessary to maintain a certain level of happiness.

Future research could explore how my results compare to estimates from different populations. Heterogeneous cross-country estimates could potentially explain the cross-country heterogeneities of income-happiness dynamics. It is also interesting to see how other ways of eliciting the same preference parameters I elicited here affect my results.

References

- Abel, Andrew B.** 1990. "Asset Prices under Habit Formation and Catching up with the Joneses." *American Economic Review*, 80(2): 38–42.
- Alvarez-Cuadrado, Francisco, Jose Maria Casado, and Jose Maria Labeaga.** 2015. "Envy and Habits: Panel Data Estimates of Interdependent Preferences." *Oxford Bulletin of Economics and Statistics*, 78(4): 443–469.
- Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew D Shapiro, and Christopher Tonetti.** 2019. "Long-Term-Care Utility and Late-in-Life Saving." *Journal of Political Economy*.
- Barsky, Robert B, F Thomas Juster, Miles S Kimball, and Matthew D Shapiro.** 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study." *The Quarterly Journal of Economics*, 112(2): 537–579.
- Benjamin, Daniel J, Kristen B Cooper, Ori Heffetz, and Miles Kimball.** 2019. "A Well-Being Snapshot in a Changing World." *American Economic Review Papers and Proceedings*, 109: 344–49.
- Benjamin, Daniel J, Ori Heffetz, Miles S Kimball, and Alex Rees-Jones.** 2012. "What Do You Think Would Make You Happier? What Do You Think You Would Choose?" *American Economic Review*, 102(5): 2083–2110.
- Benjamin, Daniel J, Ori Heffetz, Miles S Kimball, and Nichole Szembrot.** 2014. "Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference." *American Economic Review*, 104(9): 2698–2735.
- Berinsky, Adam J, Gregory A Huber, and Gabriel S Lenz.** 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon. Com's Mechanical Turk." *Political Analysis*, 20(3): 351–368.
- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer.** 2016. "Stereotypes." *The Quarterly Journal of Economics*, 131(4): 1753–1794.
- Campbell, John Y, and John H Cochrane.** 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, 107(2): 205–251.
- Cardi, Olivier, and Gernot J Müller.** 2011. "Habit Formation and Fiscal Transmission in Open Economies." *Journal of International Economics*, 85(2): 256–267.
- Carroll, Christopher D, Jody Overland, and David N Weil.** 2000. "Saving and Growth with Habit Formation." *American Economic Review*, 341–355.
- Cavallo, Alberto.** 2017. "Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers." *The American Economic Review*, 107(1): 283–303.
- Chetty, Raj, and Adam Szeidl.** 2016. "Consumption Commitments and Habit Formation." *Econometrica*, 84(2): 855–890.
- Christiano, Lawrence J, Martin S Eichenbaum, and Mathias Trabandt.** 2015. "Understanding the Great Recession." *American Economic Journal: Macroeconomics*.

economics, 7(1): 110–67.

- Clark, Andrew E.** 2016. “Adaptation and the Easterlin Paradox.” In *Advances in happiness research*. 75–94. Springer.
- Clark, Andrew E., Paul Frijters, and Michael A Shields.** 2008. “Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles.” *Journal of Economic literature*, 46(1): 95–144.
- Cochrane, John H.** 2017. “Macro-Finance.” *Review of Finance*, 21(3): 945–985.
- Constantinides, George M.** 1990. “Habit Formation: A Resolution of the Equity Premium Puzzle.” *Journal of Political Economy*, 98(3): 519–543.
- Debreu, Gerard.** 1972. “Smooth Preferences.” *Econometrica*, 40(4): 603–615.
- De Giorgi, Giacomo, Anders Frederiksen, and Luigi Pistaferri.** 2019. “Consumption Network Effects.” *The Review of Economic Studies*. rdz026.
- Diamond, Peter A., and Jerry A. Hausman.** 1994. “Contingent Valuation: Is Some Number Better Than No Number?” *The Journal of Economic Perspectives*, 8(4): 45–64.
- Di Tella, Rafael, and Robert MacCulloch.** 2008. “Gross National Happiness as an Answer to the Easterlin Paradox?” *Journal of Development Economics*, 86(1): 22–42.
- Dou, Winston Wei, Andrew W Lo, Ameya Muley, and Harald Uhlig.** 2017. “Macroeconomic Models for Monetary Policy: A Critical Review from a Finance Perspective.”
- Easterlin, Richard A.** 1973. “Does Money Buy Happiness?” *The public interest*, 30: 3.
- Easterlin, Richard A.** 1974. “Does Economic Growth Improve the Human Lot? Some Empirical Evidence.” In *Nations and households in economic growth*. 89–125. Elsevier.
- Easterlin, Richard A.** 1995. “Will Raising the Incomes of All Increase the Happiness of All?” *Journal of Economic Behavior & Organization*, 27(1): 35–47.
- Easterlin, Richard A.** 2017. “Paradox Lost?” *Review of Behavioral Economics*, 4(4): 311–339.
- Firebaugh, Glenn, and Laura Tach.** 2012. *Income, Age, and Happiness in America. Social trends in American life: Findings from the General Social Survey since 1972*. Princeton University Press.
- Flavin, Marjorie, and Shinobu Nakagawa.** 2008. “A Model of Housing in the Presence of Adjustment Costs: A Structural Interpretation of Habit Persistence.” *The American Economic Review*, 98(1): 474–495.
- Folger, Harry T.** 1926. “The Effects of Mechanical Shock on Locomotion in Amoeba Proteus.” *Journal of Morphology*, 42(2): 359–370.
- Frijters, Paul, John P Haisken-DeNew, and Michael A Shields.** 2004. “Money Does Matter! Evidence from Increasing Real Income and Life Satisfaction in East Germany Following Reunification.” *American Economic Review*, 94(3): 730–740.
- Fuhrer, Jeffrey C.** 2000. “Habit Formation in Consumption and Its Implications

- for Monetary-Policy Models.” *American Economic Review*, 367–390.
- Gardner, Jonathan, and Andrew J Oswald.** 2007. “Money and Mental Well-being: A Longitudinal Study of Medium-Sized Lottery Wins.” *Journal of health economics*, 26(1): 49–60.
- Grishchenko, Olesya V.** 2010. “Internal Vs. External Habit Formation: The Relative Importance for Asset Pricing.” *Journal of Economics and Business*, 62(3): 176–194.
- Hansen, Lars Peter, and James J Heckman.** 1996. “The Empirical Foundations of Calibration.” *Journal of Economic Perspectives*, 10(1): 87–104.
- Havranek, Tomas, Marek Rusnák, and Anna Sokolova.** 2017. “Habit Formation in Consumption: A Meta-Analysis.” *European Economic Review*.
- Houthakker, Hendrik S, and Lester D Taylor.** 1970. “Consumer Demand in the United States.”
- Jha, Saumitra, and Moses Shayo.** 2019. “Valuing Peace: The Effects of Financial Market Exposure on Votes and Political Attitudes.” *Econometrica*, 87(5): 1561–1588.
- Johnson, David, and John Ryan.** 2018. “Amazon Mechanical Turk Workers Can Provide Consistent and Economically Meaningful Data.”
- Johnston, Robert J, Kevin J Boyle, Wiktor Adamowicz, Jeff Bennett, Roy Brouwer, Trudy Ann Cameron, W Michael Hanemann, Nick Hanley, Mandy Ryan, Riccardo Scarpa, et al.** 2017. “Contemporary Guidance for Stated Preference Studies.” *Journal of the Association of Environmental and Resource Economists*, 4(2): 319–405.
- Kaiser, Casper, and Maarten Vendrik.** 2018. “Different Versions of the Easterlin Paradox: New Evidence for European Countries.”
- Kapteyn, Arie, and Federica Teppa.** 2003. “Hypothetical Intertemporal Consumption Choices.” *The Economic Journal*, 113(486): C140–C152.
- Kimball, Miles, and Robert Willis.** 2006. “Happiness and Utility.” Working Paper.
- Kimball, Miles, Claudia R Sahm, and Matthew D Shapiro.** 2008. “Imputing Risk Tolerance from Survey Responses.” *Journal of the American Statistical Association*, 103(483): 1028–1038.
- Kimball, Miles, Claudia R. Sahm, and Matthew D. Shapiro.** 2009. “Risk Preferences in the PSID: Individual Imputations and Family Covariation.” *American Economic Review*, 99(2): 363–368.
- Kimball, Miles, Fumio Ohtake, Daniel H Reck, Yoshiro Tsutsui, and Fudong Zhang.** 2015. “Diminishing Marginal Utility Revisited.”
- Kozicki, Sharon, and P.A. Tinsley.** 2002. “Dynamic Specifications in Optimizing Trend-Deviation Macro Models.” *Journal of Economic Dynamics and Control*, 26(9): 1585 – 1611. Special issue in honour of David Kendrick.
- Kuziemko, Ilyana, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva.** 2015. “How Elastic Are Preferences for Redistribution? Evi-

- dence from Randomized Survey Experiments.” *The American Economic Review*, 105(4): 1478–1508.
- Lettau, Martin, and Harald Uhlig.** 2000. “Can Habit Formation be Reconciled with Business Cycle Facts?” *Review of Economic Dynamics*, 3(1): 79–99.
- Lewbel, Arthur, Samuel Norris, Krishna Pendakur, and Xi Qu.** 2018. “Consumption Peer Effects and Utility Needs in India.” Boston College Department of Economics.
- Ljungqvist, Lars, and Harald Uhlig.** 2015. “Comment on the Campbell-Cochrane Habit Model.” *Journal of Political Economy*, 123(5): 1201–1213.
- Luttmer, Erzo F. P.** 2005. “Neighbors as Negatives: Relative Earnings and Well-Being.” *The Quarterly Journal of Economics*, 120(3): 963–1002.
- Manski, Charles F.** 2000. “Economic Analysis of Social Interactions.” *Journal of economic perspectives*, 14(3): 115–136.
- Martínez-Marquina, Alejandro, Muriel Niederle, and Emanuel Vespa.** 2019. “Failures in Contingent Reasoning: The Role of Uncertainty.” *American Economic Review*, 109(9): 3437–3474.
- Matyskova, Ludmila, Brian Rogers, Jakub Steiner, and Keh-Kuan Sun.** 2019. “Habits as Adaptations: An Experimental Study.”
- OECD.** 2013. *OECD Guidelines on Measuring Subjective Well-Being*.
- Oster, Emily, Ira Shoulson, and E Dorsey.** 2013. “Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease.” *The American Economic Review*, 103(2): 804–830.
- Perez-Truglia, Ricardo.** 2019. “The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment.” National Bureau of Economic Research.
- Ravina, Enrichetta.** 2007. “Habit Persistence and Keeping up with the Joneses: Evidence from Micro Data.”
- Rayo, Luis, and Gary S Becker.** 2007. “Habits, Peers, and Happiness: An Evolutionary Perspective.” *American Economic Review*, 97(2): 487–491.
- Sacks, Daniel W, Betsey Stevenson, and Justin Wolfers.** 2012. “The New Stylized Facts about Income and Subjective Well-Being.” *Emotion*, 12(6): 1181.
- Sahm, Claudia R.** 2007. *Stability of Risk Preference*. Division of Research & Statistics and Monetary Affairs, Federal Reserve Board.
- Smets, Frank, and Rafael Wouters.** 2007. “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach.” *The American Economic Review*, 97(3): 586–606.
- Stevenson, Betsey, and Justin Wolfers.** 2008. “Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox.” 1.
- Thurstone, Louis L.** 1931. “The Indifference Function.” *The Journal of Social Psychology*, 2(2): 139–167.
- Van Praag, Bernard M.S.** 1971. “The Welfare Function of Income in Belgium: An Empirical Investigation.” *European Economic Review*, 2(3): 337 – 369.
- Van Praag, Bernard M.S., and Paul Frijters.** 1999. “The Measurement of Wel-

fare and Well-Being; the Leyden Approach.” *Well-Being: The Foundations of Hedonic Psychology*, , ed. D. Kahneman, E. Diener and N. Schwarz, 413–432. Russel Sage Foundation.

Appendices

A Aggregation

What does the elicited preference parameters of individual respondents tell us about the preference parameters of the representative agent? This question is of particular interest because almost all current models with habit formation assume a representative agent. In this section, I prove that given its existence, the representative agent has preference parameters that are averages of the individuals' preference parameters.

To aggregate individuals, the welfare of the individuals needs to be comparable with that of each other (comparability), and the aggregate agent's welfare should represent the sum of individuals' welfare (representativeness). To formalize these two ideas, I assume that at the homogenous steady state of $\bar{C}_i = \bar{H}_i = \$5,000 \forall i$, spending an extra dollar,⁵⁴ while holding habit constant, brings the same marginal utility to every one: $u_{i,C}(\bar{C}_i, \bar{H}_i) = u_{j,C}(\bar{C}_j, \bar{H}_j) \forall i, j$.

With the comparability of individuals' utilities, the representativeness of the representative agent means that $Nu_R(C_R, H_R) = \sum_i u_i(C_i, H_i)$ when $C_R = C_i = A$ and $H_R = H_i = B \forall i$ and $\forall A, B$ in the domains of the utility functions, where N is the number of individuals in the economy. That is, when the heterogeneities in behaviors (consumption and habit) are homogenized, the representative agent is the average individual agent in terms of welfare. To see what this condition means, first note that the difference between a representative-agent model and a heterogeneous-agent model is that in the former, everyone in the economy is the same, while in the latter, each individual can be different. Imagine that everyone in the heterogeneous economy becomes the same (i.e., homogeneous in consumption, habit, and utility function, etc.), the representative agent model should behave exactly the same as the homogenized heterogeneous agents model, and hence the equality of $Nu_R(C_R, H_R) = \sum_i u_i(C_i, H_i)$. Now, allowing the individual to be heterogeneous along the dimension of utility function after the normalization of the comparability condition, this representativeness condition simply requires that the representative agent represents the individuals along the welfare dimension, after controlling for consumption and habit.

A.1 Aggregation of Habit Depreciation Rate

Even though habit depreciation rate (θ) and habit (H) are mapped one to one at each instant of time⁵⁵ for any given consumption profile, there are infinitely many pairs of them that satisfy the representative agent's habit evolution ($\dot{H}_R = \theta_R(C_R - H_R)$), individuals' habit evolutions ($\dot{H}_i = \theta_i(C_i - H_i)$), the comparability condition, and the representativeness condition. The intuition is that while H depends on the habit

⁵⁴An epsilon dollar, to be exact.

⁵⁵Before reaching a steady state.

depreciation rate, its steady-state level not. In other words, different θ leads to different H before a steady state is reached, and this difference vanishes after a steady state is reached.

To find the mapping between aggregate habit depreciation rate (θ_R) and individual habit depreciation rate (θ_i), imagine that everyone starts at the homogenous steady state and increases consumption by the same iota amount. That is, starting from $C_i = C_j = C_R = H_i = H_j = H_R \forall i, j$, increase consumption by the iota amount $\Delta C_i = \Delta C_j = \Delta C_R \forall i, j$. The changes to the utilities are

$$\begin{aligned}\Delta u_R(C_R, H_R) &= u_{R,C}(C_R, H_R) \Delta C_R + u_{R,H}(C_R, H_R) \Delta H_R \\ &= u_{R,C}(C_R, H_R) \Delta C_R + u_{R,H}(C_R, H_R) \theta_R \Delta C_R\end{aligned}$$

and

$$\begin{aligned}\Delta u_i(C_i, H_i) &= u_{i,C}(C_i, H_i) \Delta C_i + u_{i,H}(C_i, H_i) \Delta H_i \\ &= u_{i,C}(C_i, H_i) \Delta C_i + u_{i,H}(C_i, H_i) \theta_i \Delta C_i\end{aligned}$$

The representative agent representing the total welfare of individual agents implies

$$N \Delta u_R(C_R, H_R) = \sum_i \Delta u_i(C_i, H_i)$$

which implies

$$u_{R,C}(C_R, H_R) \Delta C_R + u_{R,H}(C_R, H_R) \theta_R \Delta C_R = \frac{1}{N} \sum_i [u_{i,C}(C_i, H_i) \Delta C_i + u_{i,H}(C_i, H_i) \theta_i \Delta C_i]$$

. Since $\Delta C_i = \Delta C_j = \Delta C_R$ and $u_{R,C}(C_R, H_R) = u_{i,C}(C_i, H_i)$ (comparability) $\forall i, j$,

$$\begin{aligned}\frac{u_{R,H}(C_R, H_R)}{u_{R,C}(C_R, H_R)} \theta_R &= \frac{1}{N} \sum_i \frac{u_{i,H}(C_i, H_i)}{u_{i,C}(C_i, H_i)} \theta_i \\ &= \frac{1}{N} \sum_i \frac{u_{i,H}(C_i, H_i)}{u_{i,C}(C_i, H_i)} \cdot \frac{1}{N} \sum_i \theta_i\end{aligned}$$

where the second equality holds because of the independence between slope of indifference curve and habit depreciation rate. With $u_{R,H}(C_R, H_R) = \frac{1}{N} \sum_i u_{i,H}(C_i, H_i)$ (see section A.2) and $u_{R,C}(C_R, H_R) = u_{i,C}(C_i, H_i) \forall i$, I arrive at

$$\theta_R = \frac{1}{N} \sum_i \theta_i$$

. That is, the representative agent's habit depreciation rate is the average of the individuals' habit depreciation rates.

A.2 Aggregation of Ratios of Utility Derivatives

First, I derive the relationships between utility derivatives of the representative agent and the heterogeneous agents at the baseline steady state ($\bar{C}_R = \bar{H}_R = \bar{C}_i = \bar{H}_i = \bar{C} = \bar{H} \forall i$).

Since $Nu_R(C_R, H_R) = \sum_i u_i(C_i, H_i)$ for $C_R = C_i = A$ and $H_R = H_i = B \forall i$ and $\forall A, B$ in the domains of the utility functions, utility derivatives of the representative agent is the average of the utility derivatives of the individuals:

$$u_{R,X}(C_R, H_R) = \frac{1}{N} \sum_i u_{i,X}(C_i, H_i) \xrightarrow{p} \mathbb{E}(u_{i,X}(C_i, H_i))$$

where X indicates the variable and order of the utility derivatives (e.g., C , H , CC , CH , HH).

Next, I derive the relationships between ratios of utility derivatives of the representative agent and the heterogeneous agents at the baseline steady state.

- Under the normalization of $u_{R,C}(\bar{C}, \bar{H}) = u_{i,C}(\bar{C}, \bar{H})$, the distribution of $-\frac{u_{i,H}(\bar{C}, \bar{H})}{u_{i,C}(\bar{C}, \bar{H})}$ is simply the distribution of $-u_{i,H}(\bar{C}, \bar{H})$ scaled by \bar{u} , the level of the marginal utility of consumption at the baseline ($\bar{u} = u_{R,C}(\bar{C}, \bar{H}) = u_{i,C}(\bar{C}, \bar{H})$). Thus, $\mathbb{E}(u_{i,H})$ can be calculated from

$$\mathbb{E}(u_{i,H}) = -\mathbb{E}\left(-\frac{u_{i,H}}{u_{i,C}} \cdot u_{i,C}\right) = -\mathbb{E}\left(-\frac{u_{i,H}}{u_{i,C}}\right) \cdot u_{i,C} = -\mu_{-\frac{u_{i,H}}{u_{i,C}}} \cdot \bar{u}$$

where $\mu_{-\frac{u_{i,H}}{u_{i,C}}}$ is the mean of the preference parameter $-\frac{u_{i,H}}{u_{i,C}}$ across individuals.

- The distribution of $u_{i,HH}$ is the distribution of $\frac{Hu_{i,HH}}{u_{i,H}}$ multiplied by $u_{i,H}$ and scaled by the inverse of H . Since the parameters are independent,

$$\begin{aligned} \mathbb{E}(u_{i,HH}) &= \mathbb{E}\left(\frac{Hu_{i,HH}}{u_{i,H}} \cdot u_{i,H} \cdot \frac{1}{H}\right) \\ &= \mathbb{E}\left(\frac{Hu_{i,HH}}{u_{i,H}}\right) \mathbb{E}(u_{i,H}) \frac{1}{H} \\ &= \mu_{\frac{Hu_{i,HH}}{u_{i,H}}} \cdot \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}}\right) \cdot \bar{u} \cdot \frac{1}{H} \end{aligned}$$

3. Similarly,

$$\begin{aligned}
\mathbb{E}(u_{i,CH}) &= \mathbb{E}\left(\frac{u_{i,CH}}{u_{i,HH}} \cdot u_{i,HH}\right) \\
&= \mathbb{E}\left(\frac{u_{i,CH}}{u_{i,HH}}\right) \mathbb{E}(u_{i,HH}) \\
&= \mu_{\frac{u_{i,CH}}{u_{i,HH}}} \cdot \mu_{\frac{H u_{i,HH}}{u_{i,H}}} \cdot \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}}\right) \cdot \bar{u} \cdot \frac{1}{H}
\end{aligned}$$

,

$$\begin{aligned}
\mathbb{E}(u_{i,CC}) &= \mathbb{E}\left(\frac{u_{i,CC}}{u_{i,HH}} \cdot u_{i,HH}\right) \\
&= \mathbb{E}\left(\frac{u_{i,CC}}{u_{i,HH}}\right) \mathbb{E}(u_{i,HH}) \\
&= \mu_{\frac{u_{i,CC}}{u_{i,HH}}} \cdot \mu_{\frac{H u_{i,HH}}{u_{i,H}}} \cdot \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}}\right) \cdot \bar{u} \cdot \frac{1}{H}
\end{aligned}$$

, and

$$\begin{aligned}
\mathbb{E}(u_{i,C_{others}}) &= \mathbb{E}\left(\frac{u_{i,C_{others}}}{u_{i,H}} \cdot u_{i,H}\right) \\
&= \mathbb{E}\left(\frac{u_{i,C_{others}}}{u_{i,H}}\right) \cdot \mathbb{E}(u_{i,H}) \\
&= \mu_{\frac{u_{i,C_{others}}}{u_{i,H}}} \cdot \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}}\right) \cdot \bar{u}
\end{aligned}$$

4. With these I can calculate

$$-\frac{u_{R,H}}{u_{R,C}} = -\frac{\mathbb{E}(u_{i,H})}{\mathbb{E}(u_{i,C})} = -\frac{\mathbb{E}(u_{i,H})}{\bar{u}} = \mu_{-\frac{u_{i,H}}{u_{i,C}}}$$

,

$$\frac{H u_{R,HH}}{u_{R,H}} = \frac{H \mathbb{E}(u_{i,HH})}{\mathbb{E}(u_{i,H})} = \frac{H \mu_{\frac{H u_{i,HH}}{u_{i,H}}} \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}}\right) \bar{u} \frac{1}{H}}{-\mu_{-\frac{u_{i,H}}{u_{i,C}}} \bar{u}} = \mu_{\frac{H u_{i,HH}}{u_{i,H}}}$$

,

$$\frac{u_{R,CH}}{u_{R,HH}} = \frac{\mathbb{E}(u_{i,CH})}{\mathbb{E}(u_{i,HH})} = \frac{\mu_{\frac{u_{i,CH}}{u_{i,HH}}} \mu_{\frac{H u_{i,HH}}{u_{i,H}}} \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}}\right) \bar{u} \frac{1}{H}}{\mu_{\frac{H u_{i,HH}}{u_{i,H}}} \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}}\right) \bar{u} \frac{1}{H}} = \mu_{\frac{u_{i,CH}}{u_{i,HH}}}$$

,

$$\frac{u_{R,CC}}{u_{R,HH}} = \frac{\mathbb{E}(u_{i,CC})}{\mathbb{E}(u_{i,HH})} = \frac{\mu_{\frac{u_{i,CC}}{u_{i,HH}}} \mu_{\frac{H u_{i,HH}}{u_{i,H}}} \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}} \right) \bar{u}^{\frac{1}{H}}}{\mu_{\frac{H u_{i,HH}}{u_{i,H}}} \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}} \right) \bar{u}^{\frac{1}{H}}} = \mu_{\frac{u_{i,CC}}{u_{i,HH}}}$$

, and

$$\frac{u_{R,C_{others}}}{u_{R,H}} = \frac{\mathbb{E}(u_{i,C_{others}})}{\mathbb{E}(u_{i,H})} = \frac{\mu_{\frac{u_{i,C_{others}}}{u_{i,C}}} \cdot \left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}} \right) \cdot \bar{u}}{\left(-\mu_{-\frac{u_{i,H}}{u_{i,C}}} \right) \cdot \bar{u}} = \mu_{\frac{u_{i,C_{others}}}{u_{i,H}}}$$

In summary, the representative agent's ratios of utility derivatives are averages of individuals' ratios of utility derivatives.

A.3 Aggregation of External Habit Mixture Coefficient

Imagine that everyone increases her/his consumption by the same iota amount so that the representative agent also increases her consumption by this same amount. Looking at the representativeness condition reveals that the changes in utilities satisfy

$$N [u_{R,C}(C_R, H_R) \Delta C_R + u_{R,H}(C_R, H_R) \Delta H_R] = \sum_i [u_{i,C}(C_i, H_i) \Delta C_R + u_{i,H}(C_i, H_i) \Delta H_i]$$

. Utilizing the comparability condition to get

$$N \frac{u_{R,H}(C_R, H_R)}{u_{R,C}(C_R, H_R)} \Delta H_R = \sum_i \frac{u_{i,H}(C_i, H_i)}{u_{i,C}(C_i, H_i)} \Delta H_i$$

. Since $\Delta H_R / \Delta H_i = \theta_R (1 - \omega_R) / [\theta_i (1 - \omega_i)]$, I have

$$\begin{aligned} \frac{u_{R,H}(C_R, H_R)}{u_{R,C}(C_R, H_R)} \theta_R (1 - \omega_R) &= \frac{1}{N} \sum_i \frac{u_{i,H}(C_i, H_i)}{u_{i,C}(C_i, H_i)} \theta_i (1 - \omega_i) \\ &= \left(\frac{1}{N} \sum_i \frac{u_{i,H}(C_i, H_i)}{u_{i,C}(C_i, H_i)} \right) \cdot \left(\frac{1}{N} \sum_i \theta_i \right) \cdot \left(\frac{1}{N} \sum_i (1 - \omega_i) \right) \end{aligned}$$

where the second equality because of the independence between the preference parameters. Finally, by $\frac{u_{R,H}(C_R, H_R)}{u_{R,C}(C_R, H_R)} = \frac{1}{N} \sum_i \frac{u_{i,H}(C_i, H_i)}{u_{i,C}(C_i, H_i)}$ and $\theta_R = \frac{1}{N} \sum_i \theta_i$,

$$\omega_R = \frac{1}{N} \sum_i \omega_i$$

B Equivalence of Linear and Nonlinear Habit Evolutions

In this section, I show that the model with the linear habit evolution (Model L below) and models with nonlinear habit evolutions (Model N below) are observationally equivalent. The intuition is that linear habit (H) and nonlinear habit (\mathcal{H}) are two measurements of the same thing - habit. Just like using foot and meter to measure length does not change the effect of length, using H and \mathcal{H} to measure habit will not change the effect of habit (on optimal choices).⁵⁶

- Model L :

$$\begin{aligned} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(C, H) dt \\ s.t. \quad \dot{H} = \theta(C - H) \end{aligned}$$

- Model N :

$$\begin{aligned} \mathbb{E}_0 \int_0^\infty e^{-\rho t} v(C, \mathcal{H}) dt \\ s.t. \quad \dot{\mathcal{H}} = f(C, \mathcal{H}) \end{aligned}$$

where f can be a nonlinear function of C and \mathcal{H} .

Note that $H_t = h(C_0, H_0, t)$ if $C_t = C_{t+1}$ for all $t \geq 0$ (subscripts index time). Similarly, $\mathcal{H}_t = k(C_0, \mathcal{H}_0, t)$ if $C_t = C_{t+1}$ for all $t \geq 0$. That is, when the consumption path does not change, H_t and \mathcal{H}_t are functions of only time while C_0 , H_0 , and \mathcal{H}_0 are their parameters.

Definition 1. Monotonocities of two functions are entangled with respect to a variable if 1) the two functions share this variable as an argument and 2) when one function is monotonic in the argument, the other function is also monotonic in the argument.

Because H and \mathcal{H} are two measurements of one fundamental - habit, they change at the same time (though in potentially different ways) when habit changes and stop changing when habit stops changing. By Definition 1, their monotonocities⁵⁷ are entangled with respect to time.

Definition 2. Two models are observationally equivalent if they lead to the same set of optimal choices.

⁵⁶For example, the geometric habit evolution, as in Kozicki and Tinsley (2002), synchronizes with the linear habit evolution of Model L .

⁵⁷That is, H' and \mathcal{H}' can equal to 0 but they will not change sign: i.e., $H' \cdot \mathcal{H}'$ does not change sign around region that $H' \cdot \mathcal{H}' = 0$.

Proposition 11. *Model L and Model N are observationally equivalent if monotonicities of H and H̄ are entangled with respect to time.*

Proof. Since H and H̄ are entangled monotonically with respect to time, without loss of generality, suppose that H and H̄ are monotonic from period 0 to period T and flat afterward (i.e., remain at constant levels), say at levels H̄ and H̄. Suppose also consumption changes at period 0 and stays at that level: C_t = C_{t+1} ≠ C_{-ε} for all t ≥ 0 and ε > 0. As a result, H_t = a(t|C₀, H₀) and H̄_t = b(t|C₀, H̄₀) where a(·) and b(·) are monotonic functions of t for 0 ≤ t ≤ T and flat for t > T, where T is the instant when habit reaches its new steady state.

Since

$$H_t = b(t|C_0, H_0) = b(a^{-1}(a(t|C_0, H_0))|C_0, H_0) = b(a^{-1}(H_t)|C_0, H_0)$$

for 0 ≤ t ≤ T and

$$H_t = \frac{\bar{H}}{\bar{H}} H_t$$

for t > T, there always exists an injective function G that maps H_t into H̄_t:

$$H_t = G(H_t) = \begin{cases} b(a^{-1}(H_t)|C_0, H_0) & 0 \leq t \leq T \\ \frac{\bar{H}}{\bar{H}} H_t & t > T \end{cases}$$

. For other patterns of monotonicities of the two functions (e.g. flat to monotonic to flat to monotonic, etc.), the function G can be derived analogously.

Since

$$v(C(t), H(t)) = v(C(t), G(H(t))) \equiv u(C(t), H(t))$$

, Model N gives the same utility as Model L for any consumption path that is constant for t ≥ 0.

When consumption path is not constant over time, the above remains true. To see this, start from the period when consumption is changed for the last time and apply the above logic to get the same utility from the two models starting from that period onwards. Then go back to the period when consumption is changed for the second to last time and apply the above logic. Same utility results again for the two models. Continue this process until the first period of interest.

Since the utilities from the two models are the same, the consumption choices generated from these two models coincide. Suppose not. That is, the two models generate different optimal consumption paths, {C_L^{*}} ≠ {C_N^{*}} for at least one instant, where

$$\{C_L^*\} = \arg \max_{\{C\}} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(C, H) dt \equiv \arg \max_{\{C_L\}} U(\{C_L\}, H_0)$$

and

$$\{C_N^*\} = \arg \max_{\{C\}} \mathbb{E}_0 \int_0^\infty e^{-\rho t} v(C, H) dt \equiv \arg \max_{\{C_N\}} V(\{C_N\}, H_0)$$

. If $U(\{C_L^*\}, H_0) \neq V(\{C_N^*\}, \mathcal{H}_0)$, at least one of the two consumption paths is not the optimal solution, contradicting that both of them are optimal solutions. If $U(\{C_L^*\}, H_0) = V(\{C_N^*\}, \mathcal{H})$, then the consumption path $\{C_L^*\}$ will also be a solution to the Model N while $\{C_N^*\}$ also a solution to Model L . Therefore, $\{C_H^*\}$ and $\{C_{\mathcal{H}}^*\}$ are both solutions to the two models. In other words, the two models have the same set of solutions. Thus, by Definition 2, the two models are observationally equivalent. \square

Since the monotonicities of H and \mathcal{H} are entangled with respect to time, Model L and Model N are observationally equivalent.

C Elicitation of Time Discount Rate

Since the time discount rate is of only indirect interest⁵⁸, it is relegated to here.

To elicit the time discount rate, I increase spending in the next year and the year after next year. The resulting survey question has monthly spending graphs as in Figure 22.

Proposition 12. *Under exact elicitation, $\rho > -\ln \frac{\Delta e}{\Delta f}$ if the respondent chooses Universe One over Universe Two for a better future experience in the time discount rate question.⁵⁹*

Estimation pins down a value of about 0.15 for the time discount rate (Table 24).

Table 24: Estimate of Time Discount Rate

	MAP	Mean	Median	95% HPDI
ρ	0.15	0.14	0.14	[0.04, 0.25]

D Proofs

D.1 Lemmas

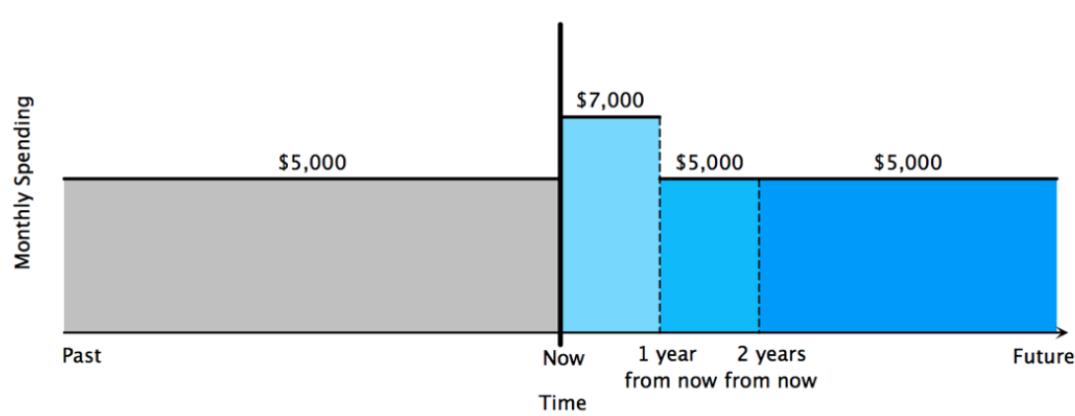
I prove three lemmas that will be used in proving propositions that have not been proved thus far.

Lemma 1. *For $a, b, c \in \mathbb{R}$, if $a(a + b) > 0$, then $a(a + cb) > 0$ as long as $0 \leq c \leq 1$.*

⁵⁸That is, helping the estimation of other preference parameters.

⁵⁹ $\Delta e = \$2,000$ and $\Delta f = \$2,200$ in the example of Figure 22.

Universe One



Universe Two

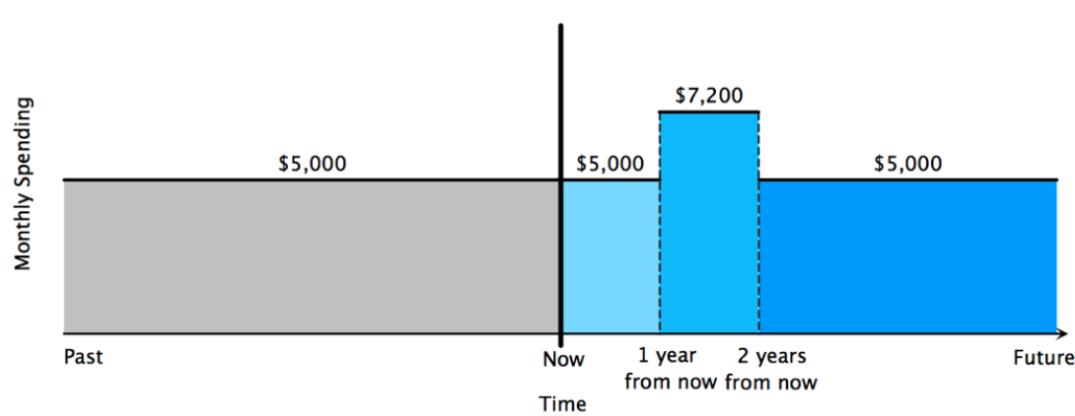


Figure 22: Monthly Spending Graphs - Time Discount Rate

Proof. $a(a+b) > 0$ is equivalent to $a+b < 0$ if $a < 0$ and $a+b > 0$ if $a > 0$.

Suppose $a < 0$ and $a+b < 0$. Note that $a+cb = a+b + (c-1)b$. If $b \geq 0$, then $(c-1)b \leq 0$ and thus $a+cb \leq a+b < 0$. If $b < 0$, then by $a < 0$ and $c \geq 0$, $a+cb < 0$. Therefore $a(a+cb) > 0$.

Suppose $a > 0$ and $a+b > 0$. If $b \leq 0$, then $(c-1)b \geq 0$ and therefore $a+cb \geq a+b > 0$. If $b > 0$, then by $a > 0$ and $c \geq 0$, $a+cb > 0$. Therefore $a(a+cb) > 0$. \square

Lemma 2. For $\Delta e, \Delta f, M, \theta \in \mathbb{R}^+$, if $M - \Delta f - \Delta e \geq 0$, $\frac{\sum_{s=0}^n (\Delta e)^s (\Delta f)^{n-s}}{M^n}$ is non-increasing in $n \in \mathbb{N}^+$.

Proof. $\frac{\sum_{s=0}^n (\Delta e)^s (\Delta f)^{n-s}}{M^n}$ is non-increasing in $n \in \mathbb{N}^+$ if $\frac{\sum_{s=0}^n (\Delta e)^s (\Delta f)^{n-s}}{M^n} - \frac{\sum_{s=0}^{n+1} (\Delta e)^s (\Delta f)^{n+1-s}}{M^{n+1}} \geq 0 \forall n \in \mathbb{N}^+$.

Now

$$\begin{aligned}
& \frac{\sum_{s=0}^n (\Delta e)^s (\Delta f)^{n-s}}{M^n} - \frac{\sum_{s=0}^{n+1} (\Delta e)^s (\Delta f)^{n+1-s}}{M^{n+1}} \\
&= \frac{\sum_{s=0}^n (\Delta e)^s (\Delta f)^{n-s}}{M^n} - \frac{\sum_{s=0}^n (\Delta e)^s (\Delta f)^{n-s} + \frac{(\Delta e)^{n+1}}{\Delta f} \Delta f}{M^{n+1}} \\
&= \frac{\sum_{s=0}^n (\Delta e)^s (\Delta f)^{n-s}}{M^n} \left(1 - \frac{\Delta f}{M} \right) - \frac{(\Delta e)^{n+1}}{M^{n+1}} \\
&= \frac{1}{M^{n+1}} \left[\sum_{s=0}^n (\Delta e)^s (\Delta f)^{n-s} (M - \Delta f) - (\Delta e)^{n+1} \right] \\
&= \frac{(\Delta e)^n}{M^{n+1}} \left[\sum_{s=0}^n \left(\frac{\Delta f}{\Delta e} \right)^{n-s} (M - \Delta f) - \Delta e \right] \\
&= \frac{(\Delta e)^n (M - \Delta f)}{M^{n+1}} \left[\sum_{s=0}^n \left(\frac{\Delta f}{\Delta e} \right)^{n-s} - \frac{\Delta e}{M - \Delta f} \right] \\
&= \frac{(\Delta e)^n (M - \Delta f)}{M^{n+1}} \left[\sum_{s=0}^{n-1} \left(\frac{\Delta f}{\Delta e} \right)^{n-s} + 1 - \frac{\Delta e}{M - \Delta f} \right] \\
&= \frac{(\Delta e)^n (M - \Delta f)}{M^{n+1}} \left[\sum_{s=0}^{n-1} \left(\frac{\Delta f}{\Delta e} \right)^{n-s} + \frac{M - \Delta f - \Delta e}{M - \Delta f} \right] \\
&\geq \frac{(\Delta e)^n (M - \Delta f)}{M^{n+1}} \left[\sum_{s=0}^{n-1} \left(\frac{\Delta f}{\Delta e} \right)^{n-s} \right] \\
&\geq 0
\end{aligned}$$

where the first inequality holds because $M - \Delta f - \Delta e \geq 0$ and $\Delta e > 0$. \square

Lemma 3. For $\Delta e, \Delta f, M, \theta \in \mathbb{R}^+$, if $M - \Delta f - \Delta e \geq 0$ and utility function is analytic and $u_H < 0$, then

$$u_H e^{-\theta t} + \frac{1}{2} u_{HH} (\Delta f + \Delta e) (e^{-\theta t})^2 + \cdots + \frac{1}{n!} \frac{\partial^n u}{\partial H^n} \left(\sum_{s=0}^{n-1} (\Delta e)^s (\Delta f)^{n-s} \right) (e^{-\theta t})^n + \cdots < 0$$

. By $u_H < 0$ and analyticity of the utility function,

$$u(C, H + M) - u(C, H) = u_H M + \frac{1}{2} u_{HH} M^2 + \cdots + \frac{1}{n!} \frac{\partial^n u}{\partial H^n} M^n + \cdots < 0$$

for $M > 0$. Since $0 \leq \frac{(\Delta f + \Delta e)}{M} e^{-\theta t} \leq 1$ for $t \in \mathbb{N}^+$, by Lemma 1

$$u_H M + \frac{\Delta f + \Delta e}{M} e^{-\theta t} \left[\frac{1}{2} u_{HH} M^2 + \cdots + \frac{1}{n!} \frac{\partial^n u}{\partial H^n} M^n + \cdots \right] < 0$$

. By Lemma 2,

$$\frac{\sum_{s=0}^2 (\Delta e)^s (\Delta f)^{2-s}}{M^2} (e^{-\theta t})^2 = \frac{\sum_{s=0}^2 (\Delta e)^s (\Delta f)^{2-s}}{\frac{(\Delta f + \Delta e)}{M} e^{-\theta t}} e^{-\theta t} \leq 1$$

. Apply Lemma 1 again to get

$$\begin{aligned} & u_H M + \frac{1}{2} u_{HH} M^2 \frac{\Delta f + \Delta e}{M} e^{-\theta t} \\ & + \frac{\sum_{s=0}^2 (\Delta e)^s (\Delta f)^{2-s}}{\frac{(\Delta f + \Delta e)}{M}} e^{-\theta t} \frac{(\Delta f + \Delta e)}{M} e^{-\theta t} \left[\frac{1}{3!} u_{HHH} M^3 + \cdots + \frac{1}{n!} \frac{\partial^n u}{\partial H^n} M^n + \cdots \right] \\ & = u_H M + \frac{1}{2} u_{HH} M^2 \frac{\Delta f + \Delta e}{M} e^{-\theta t} \\ & + \frac{\sum_{s=0}^2 (\Delta e)^s (\Delta f)^{2-s}}{M^2} (e^{-\theta t})^2 \left[\frac{1}{3!} u_{HHH} M^3 + \cdots + \frac{1}{n!} \frac{\partial^n u}{\partial H^n} M^n + \cdots \right] \\ & < 0 \end{aligned}$$

Repeating this process for all other $n \in \mathbb{N}^+$ to get

$$u_H e^{-\theta t} + \frac{1}{2} u_{HH} (e^{-\theta t})^2 (\Delta f + \Delta e) + \cdots + \frac{1}{n!} \frac{\partial^n u}{\partial H^n} \left(\sum_{s=0}^{n-1} (\Delta e)^s (\Delta f)^{n-s} \right) (e^{-\theta t})^n + \cdots < 0$$

. \square

D.2 Proof of Proposition 1

Proof. That θ is habit depreciation rate implies $\theta \in \mathbb{R}^+$. Taking $M = \$5000$, then $M - \Delta C_{U1} - (1 - e^{-\theta}) \Delta C_{U2} > 0$ in all the questions for habit depreciation rate.⁶⁰

A respondent preferring Universe One for a better future experience implies

$$\begin{aligned}
& \int_0^\infty e^{-\rho t} \left[u_H e^{-\theta t} \Delta C_{U1} + \frac{1}{2} u_{HH} (e^{-\theta t} \Delta C_{U1})^2 + \dots \right] dt \\
& - \int_0^\infty e^{-\rho t} \left[u_H e^{-\theta t} (1 - e^{-\theta}) \Delta C_{U2} + \frac{1}{2} u_{HH} (e^{-\theta t} (1 - e^{-\theta}) \Delta C_{U2})^2 + \dots \right] dt \\
& = \int_0^\infty e^{-\rho t} \left\{ u_H e^{-\theta t} [\Delta C_{U1} - (1 - e^{-\theta}) \Delta C_{U2}] \right. \\
& \quad \left. + \frac{1}{2} u_{HH} (e^{-\theta t})^2 [(\Delta C_{U1})^2 - ((1 - e^{-\theta}) \Delta C_{U2})^2] + \dots \right\} dt \\
& = [\Delta C_{U1} - (1 - e^{-\theta}) \Delta C_{U2}] \int_0^\infty e^{-\rho t} \left\{ u_H e^{-\theta t} \right. \\
& \quad \left. + \frac{1}{2} u_{HH} (e^{-\theta t})^2 [\Delta C_{U1} + (1 - e^{-\theta}) \Delta C_{U2}] + \dots \right\} dt \\
& > 0
\end{aligned}$$

which, by Lemma 3, implies

$$\Delta C_{U1} - (1 - e^{-\theta}) \Delta C_{U2} < 0$$

or

$$\theta > -\ln \left(1 - \frac{\Delta C_{U1}}{\Delta C_{U2}} \right)$$

□

D.3 Proof of Proposition 2

Proof.

$$\frac{u_{CH}}{u_{HH}} \frac{u_H}{u_C} = \frac{-\gamma v''}{\gamma^2 v''} \frac{-\gamma v'}{v'} = 1$$

, and

$$\frac{u_{CH}}{u_{CC}} \frac{u_C}{u_H} = \frac{-\gamma v''}{v''} \frac{v'}{-\gamma v'} = 1$$

□

⁶⁰See Table 26 for all the values of ΔC_{U1} and ΔC_{U2} used in survey.

D.4 Proof of Proposition 3

Proof.

$$\frac{Hu_H u_{CH}}{u_C (u_H + Hu_{HH})} = \frac{H \left(-\gamma \frac{C}{H^{\gamma+1}} v' \right) \left(-\gamma \frac{1}{H^{\gamma+1}} v' - \gamma \frac{C}{H^{2\gamma+1}} v'' \right)}{\frac{1}{H^\gamma} v' \left\{ -\gamma \frac{C}{H^{\gamma+1}} v' + H \gamma \frac{C}{H^2} \left[(\gamma+1) \frac{1}{H^\gamma} v' + \gamma \frac{C}{H^{2\gamma}} v'' \right] \right\}} \\ = 1$$

, and

$$\frac{Cu_C u_{CH}}{u_H (u_C + Cu_{CC})} = \frac{C \frac{1}{H^\gamma} v' \left(-\gamma \frac{1}{H^{\gamma+1}} v' - \gamma \frac{C}{H^{2\gamma+1}} v'' \right)}{\left(-\gamma \frac{C}{H^{\gamma+1}} v' \right) \left(\frac{1}{H^\gamma} v' + C \frac{1}{H^{2\gamma}} v'' \right)} \\ = 1$$

□

D.5 Proof of Proposition 4

Proof. Under second-order elicitation, a respondent preferring Universe One for a better future experience implies

$$\begin{aligned} & \frac{1}{\rho} \left\{ u_C \Delta f + \frac{1}{\rho + \theta} u_H (\rho \Delta e + \theta \Delta f) + \frac{1}{2} \left[u_{CC} (\Delta f)^2 + 2 \frac{1}{\rho + \theta} u_{CH} \Delta f (\rho \Delta e + \theta \Delta f) \right. \right. \\ & + \frac{1}{(\rho + \theta)(\rho + 2\theta)} u_{HH} (\rho(\rho + \theta)(\Delta e)^2 + 2\rho\theta\Delta e\Delta f + 2\theta^2(\Delta f)^2) \left. \right] \left\} \right. \\ & - \frac{1}{\rho} \left\{ u_C (-\Delta f) + \frac{1}{\rho + \theta} u_H (\rho(-\Delta e) + \theta(-\Delta f)) + \frac{1}{2} \left[u_{CC} (\Delta f)^2 \right. \right. \\ & + 2 \frac{1}{\rho + \theta} u_{CH} \Delta f (\rho \Delta e + \theta \Delta f) + \frac{1}{(\rho + \theta)(\rho + 2\theta)} u_{HH} (\rho(\rho + \theta)(\Delta e)^2 \\ & \left. \left. + 2\rho\theta\Delta e\Delta f + 2\theta^2(\Delta f)^2 \right] \right\} \\ & = \frac{2}{\rho} \left[u_C \Delta f + \frac{1}{\rho + \theta} u_H (\rho \Delta e + \theta \Delta f) \right] \\ & > 0 \end{aligned}$$

which, by $u_C > 0$, implies

$$-\frac{u_H}{u_C} < \frac{(\rho + \theta) \Delta f}{\rho \Delta e + \theta \Delta f}$$

□

D.6 Proof of Proposition 5

Proof. Under second-order elicitation, a respondent preferring Universe One for a better future experience implies

$$\begin{aligned} & \frac{1}{2} \left[\frac{1}{\rho + \theta} u_H (-\Delta f) + \frac{1}{2\rho + 2\theta} u_{HH} (\Delta f)^2 \right] + \frac{1}{2} \left[\frac{1}{\rho + \theta} u_H \Delta e + \frac{1}{2\rho + 2\theta} u_{HH} (\Delta e)^2 \right] \\ &= \frac{1}{2} \left\{ \frac{1}{\rho + \theta} u_H (\Delta e - \Delta f) + \frac{1}{2\rho + 2\theta} u_{HH} [(\Delta f)^2 + (\Delta e)^2] \right\} \\ &> 0 \end{aligned}$$

which, by $u_H < 0$,⁶¹ implies

$$\frac{Hu_{HH}}{u_H} < \frac{2(\rho + 2\theta)}{\rho + \theta} \frac{\Delta f/\Delta e - 1}{(\Delta f/\Delta e)^2 + 1} \frac{H}{\Delta e}$$

.

□

⁶¹This sign is elicited in the existence of internal habit formation question.

D.7 Proof of Proposition 6

Proof. Under third-order elicitation, a respondent preferring Universe One for a better future experience implies

$$\begin{aligned}
& \frac{1}{2} \frac{1}{\rho} \left\{ u_C \Delta f + \frac{1}{\rho + \theta} u_H (\rho \Delta e + \theta \Delta f) + \frac{1}{2} \left[u_{CC} (\Delta f)^2 + 2 \frac{1}{\rho + \theta} u_{CH} \Delta f (\rho \Delta e + \theta \Delta f) \right. \right. \\
& \quad \left. \left. + \frac{1}{(\rho + \theta)(\rho + 2\theta)} u_{HH} (\rho(\rho + \theta)(\Delta e)^2 + 2\rho\theta\Delta e\Delta f + 2\theta^2(\Delta f)^2) \right] \right\} \\
& + \frac{1}{2} \frac{1}{\rho} \left\{ -u_C \Delta f - \frac{1}{\rho + \theta} u_H (\rho \Delta e + \theta \Delta f) + \frac{1}{2} \left[u_{CC} (\Delta f)^2 + 2 \frac{1}{\rho + \theta} u_{CH} \Delta f (\rho \Delta e + \theta \Delta f) \right. \right. \\
& \quad \left. \left. + \frac{1}{(\rho + \theta)(\rho + 2\theta)} u_{HH} (\rho(\rho + \theta)(\Delta e)^2 + 2\rho\theta\Delta e\Delta f + 2\theta^2(\Delta f)^2) \right] \right\} \\
& - \frac{1}{2} \frac{1}{\rho} \left\{ \left[u_C + \frac{\theta}{\rho + \theta} u_H \right] \Delta f + \frac{1}{2} \left[u_{CC} + 2u_{CH} \frac{\theta}{\rho + \theta} + u_{HH} \frac{2\theta^2}{(\rho + \theta)(\rho + 2\theta)} \right] (\Delta f)^2 \right\} \\
& - \frac{1}{2} \frac{1}{\rho} \left\{ \left[u_C + \frac{\theta}{\rho + \theta} u_H \right] (-\Delta f) + \frac{1}{2} \left[u_{CC} + 2u_{CH} \frac{\theta}{\rho + \theta} + u_{HH} \frac{2\theta^2}{(\rho + \theta)(\rho + 2\theta)} \right] (\Delta f)^2 \right\} \\
& = \frac{1}{2} \frac{1}{\rho} \left[u_{CC} (\Delta f)^2 + 2 \frac{1}{\rho + \theta} u_{CH} \Delta f (\rho \Delta e + \theta \Delta f) \right. \\
& \quad \left. + \frac{1}{(\rho + \theta)(\rho + 2\theta)} u_{HH} (\rho(\rho + \theta)(\Delta e)^2 + 2\rho\theta\Delta e\Delta f + 2\theta^2(\Delta f)^2) \right] \\
& \quad - \frac{1}{2} \frac{1}{\rho} \left[u_{CC} + 2u_{CH} \frac{\theta}{\rho + \theta} + u_{HH} \frac{2\theta^2}{(\rho + \theta)(\rho + 2\theta)} \right] (\Delta f)^2 \\
& = \frac{1}{2} \frac{1}{\rho} \left[2u_{CH} \frac{\rho}{\rho + \theta} \Delta f \Delta e + \frac{1}{(\rho + \theta)(\rho + 2\theta)} u_{HH} (\rho(\rho + \theta)(\Delta e)^2 + 2\rho\theta\Delta e\Delta f) \right] \\
& > 0
\end{aligned}$$

which, by $u_{HH} < 0$,⁶² implies

$$\frac{u_{CH}}{u_{HH}} < -\frac{(\rho + \theta) \Delta e + 2\theta \Delta f}{2(\rho + 2\theta) \Delta f}$$

.

□

⁶²This sign is elicited in the Hu_{HH}/u_H question.

D.8 Proof of Proposition 7

Proof. Under third-order elicitation, a respondent preferring Universe One for a better future experience implies

$$\begin{aligned}
& \frac{1}{2} \frac{1}{\rho} \left\{ \left[u_C + \frac{\theta}{\rho + \theta} u_H \right] \Delta f + \frac{1}{2} \left[u_{CC} + 2u_{CH} \frac{\theta}{\rho + \theta} + u_{HH} \frac{2\theta^2}{(\rho + \theta)(\rho + 2\theta)} \right] (\Delta f)^2 \right\} \\
& + \frac{1}{2} \frac{1}{\rho} \left\{ \left[u_C + \frac{\theta}{\rho + \theta} u_H \right] (-\Delta f) + \frac{1}{2} \left[u_{CC} + 2u_{CH} \frac{\theta}{\rho + \theta} + u_{HH} \frac{2\theta^2}{(\rho + \theta)(\rho + 2\theta)} \right] (\Delta f)^2 \right\} \\
& - \frac{1}{2} \left[\frac{1}{\rho + \theta} u_H \Delta e + \frac{1}{2} \frac{1}{\rho + 2\theta} u_{HH} (\Delta e)^2 \right] - \frac{1}{2} \left[\frac{1}{\rho + \theta} u_H (-\Delta e) + \frac{1}{2} \frac{1}{\rho + 2\theta} u_{HH} (\Delta e)^2 \right] \\
& = \frac{1}{2} \frac{1}{\rho} \left[u_{CC} + 2u_{CH} \frac{\theta}{\rho + \theta} + u_{HH} \frac{2\theta^2}{(\rho + \theta)(\rho + 2\theta)} \right] (\Delta f)^2 - \frac{1}{2} \frac{1}{\rho + 2\theta} u_{HH} (\Delta e)^2 \\
& > 0
\end{aligned}$$

which, by $u_{HH} < 0$,⁶³ implies

$$\frac{u_{CC}}{u_{HH}} < \frac{\rho}{\rho + 2\theta} \left(\frac{\Delta e}{\Delta f} \right)^2 - \frac{2\theta}{\rho + \theta} \frac{u_{CH}}{u_{HH}} - \frac{2\theta^2}{(\rho + \theta)(\rho + 2\theta)}$$

□

D.9 Proof of Proposition 8

Proof. Choosing Universe One over Universe Two implies

$$u(C, H(\Delta C > 0, \Delta C_{others} = 0)) > u(C, H(\Delta C = 0, \Delta C_{others} > 0))$$

. Subtracting the baseline utility from both sides to get

$$\begin{aligned}
& u(C, H(\Delta C = 0, \Delta C_{others} > 0)) - u(C, H(\Delta C = 0, \Delta C_{others} = 0)) \\
& < u(C, H(\Delta C > 0, \Delta C_{others} = 0)) - u(C, H(\Delta C = 0, \Delta C_{others} = 0)) \\
& < 0
\end{aligned}$$

where the last inequality because of the existence of internal habit formation. □

D.10 Proof of Proposition 9

Proof. Take $M = \$5000$, then $M - (1 - \omega) \Delta C - \omega \Delta C_{others} > 0$ in all the questions for external habit.

⁶³This sign is elicited in the $H u_{HH}/u_H$ question.

A respondent preferring Universe One for a better future experience implies

$$\begin{aligned}
& \int_0^\infty e^{-\rho t} \left[u_H e^{-\theta t} ((1-\omega) \Delta C) + \frac{1}{2} u_{HH} (e^{-\theta t} ((1-\omega) \Delta C))^2 + \dots \right] dt \\
& - \int_0^\infty e^{-\rho t} \left[u_H e^{-\theta t} (\omega \Delta C_{others}) + \frac{1}{2} u_{HH} (e^{-\theta t} (\omega \Delta C_{others}))^2 + \dots \right] dt \\
& = [(1-\omega) \Delta C - \omega \Delta C_{others}] \int_0^\infty e^{-\rho t} \left[u_H e^{-\theta t} \right. \\
& \quad \left. + \frac{1}{2} u_{HH} (e^{-\theta t})^2 (((1-\omega) \Delta C + \omega \Delta C_{others})) + \dots \right] dt \\
& > 0
\end{aligned}$$

which, by Lemma 3, implies

$$(1-\omega) \Delta C - \omega \Delta C_{others} < 0$$

or

$$\omega > \frac{\Delta C}{\Delta C + \Delta C_{others}}$$

□

D.11 Proof of Proposition 10

Proof. Under first-order elicitation, a respondent preferring Universe One for a better future experience implies

$$\frac{1}{\rho} \left[u_{C_{others}} + \frac{\theta \omega}{\rho + \theta} u_H \right] \Delta C_{others}^{U1} - \frac{\omega}{\rho + \theta} u_H \Delta C_{others}^{U2} > 0$$

which, by $u_H < 0$,⁶⁴ implies

$$\frac{u_{C_{others}}}{u_H} < \frac{\omega}{\rho + \theta} \left(\rho \frac{\Delta C_{others}^{U2}}{\Delta C_{others}^{U1}} - \theta \right)$$

□

D.12 Proof of Proposition 12

Proof. Let

$$\begin{aligned}
Q(a) \equiv & \int_0^1 e^{-\rho t} \left\{ u_C a + \frac{1}{2} u_{CC} a^2 + u_{CH} (\Delta c (1 - e^{-\theta t}) a) + \dots \right. \\
& + \left. \left[u_H (1 - e^{-\theta t}) a + \frac{1}{2} u_{HH} ((1 - e^{-\theta t}) a)^2 + \dots \right] \right\} dt \\
& + \int_1^\infty e^{-\rho t} \left[u_H e^{-\theta t} (1 - e^{-\theta}) a + \frac{1}{2} u_{HH} (e^{-\theta t} (1 - e^{-\theta}) a)^2 + \dots \right] dt
\end{aligned}$$

⁶⁴This sign is elicited in the existence of habit formation question.

Under exact elicitation, a respondent preferring Universe One for a better future experience implies

$$\begin{aligned} Q(\Delta e) - e^{-\rho}Q(\Delta f) &= \Delta e \frac{Q(\Delta e)}{\Delta e} - e^{-\rho}\Delta f \frac{Q(\Delta f)}{\Delta f} \\ &> (\Delta e - e^{-\rho}\Delta f) \frac{Q(\Delta e)}{\Delta e} \\ &> 0 \end{aligned}$$

where the first inequality holds because of diminishing marginal utility $\frac{Q(\Delta e)}{\Delta e} > \frac{Q(\Delta f)}{\Delta f} > 0$ for $\Delta f > \Delta e > 0$. This implies

$$\Delta e - e^{-\rho}\Delta f > 0$$

or

$$\rho > -\ln \frac{\Delta e}{\Delta f}$$

□

E Response Frequency

See Table 25.

Table 25: Response Frequency (Percentage) for Preference Parameters Identifiable to Scale

Question	Wave	Choice					
		1	2	3	4	5	6
Habit Depreciation Rate	1	28	10	16	11	7	28
	2	30	10	14	11	6	30
Slope of Indifference Curve	1	32	3	7	7	12	39
	2	31	4	6	2	20	36
Hu_{HH}/u_H	1	14	4	8	10	3	60
	2	25	2	6	5	1	61
u_{CH}/u_{HH}	1	7	4	11	10	27	41
	2	9	3	10	9	34	36
u_{CC}/u_{HH}	1	24	30	10	9	5	22
	2	24	19	8	10	9	30
External Habit Mixture Coefficient	1	46	9	13	8	4	18
	2	46	4	7	16	5	23
$u_{C_{others}}/u_H$	1	26	18	10	8	2	36
	2	24	19	8	6	3	40
Time Discount Rate	1	34	4	7	7	4	43
	2	39	3	5	5	4	45

F Survey

F.1 Instructions

Instruction 1/2

The survey starts with an instruction that specifies the hypothetical situation discussed in section 2.2. A set of practice questions follows to test the respondents' understanding of the hypothetical situation:

With no inflation and prices of everything staying the same, if you can buy 3 bananas with one dollar in the last year, how many bananas can you buy with one dollar in the next year?

- 5
- 3
- 1
- No idea

If you rent the durable goods you consume, select any of the following that you own (that is, not rent):

- Residence
- Car
- Furniture
- I do not own any of the above
- No idea

Instruction 2/2

After passing the practice questions, the respondents are presented with an instruction on how to read monthly spending graphs.

In this survey, you'll compare your experience in several universes that are identical except that your monthly spending differs.

- Monthly spending refers to the total amount of money you spend, rather than earn, in each month.
- You will be asked to find out in which universe you will have (had) a better experience given how much you spend (spent).
- 'Better' means more satisfying.
- You can afford the monthly spending specified in the questions.

The difference of your monthly spending between the universes is detailed in monthly spending graphs, like the one below.

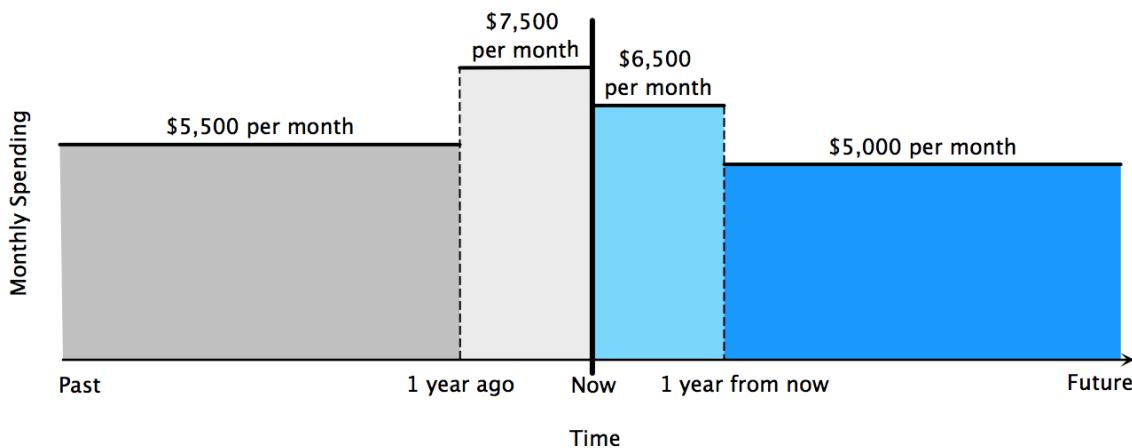


Figure 23: Instruction - Monthly Spending Graph 1

Now let's learn to read a monthly spending graph.

The first element of a monthly spending graph is the timeline, with past on the left, now in the middle, future on the right. A thick vertical line representing now separates the past from the future.

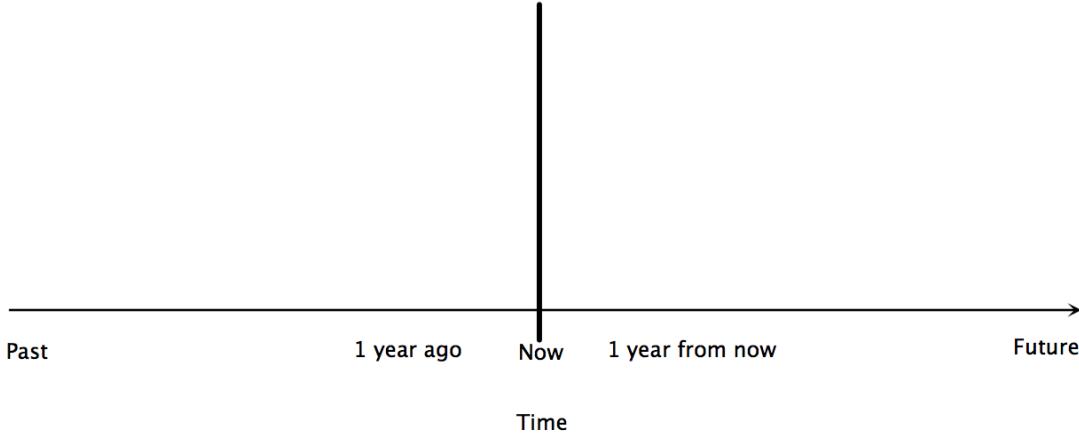


Figure 24: Instruction - Timeline

To fix the idea, the ‘Past’ means as far back in the past as you can remember and the ‘Future’ as far in the future as you can imagine. If easier, think of the ‘Past’ as the past 30 years and the ‘Future’ as the next 30 years.

The second element of a monthly spending graph is the bars above the timeline.

- The height of bars represents the level of monthly spending (again, not income) in time frames covered by the bars.
- The exact level of monthly spending is labeled on top of the corresponding bar. The words ‘per month’ are saved for space consideration from now on, but you should always remember that the numbers are per month spending.
- The bars are colored differently to help you distinguish different time frames.

For example, if the following monthly spending graph describes your monthly spending,

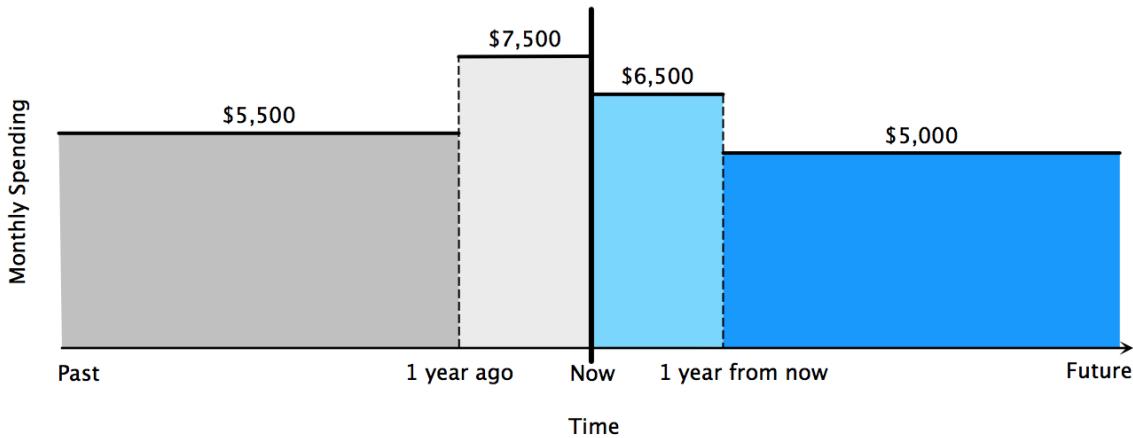


Figure 25: Instruction - Monthly Spending Graph 2

you spent/spend

- \$5,500 per month in the 'past' until '1 year ago';
- \$7,500 per month from '1 year ago' until 'now' (or in the 'past year');
- \$6,500 per month from 'now' to '1 year from now' (or in the 'next year');
- \$5,000 per month from '1 year from now' onward.

To highlight the difference of monthly spending, the time frames as in the above example are sometimes collapsed into three or two timeframes. For instance, if in Universe One your monthly spending graph is

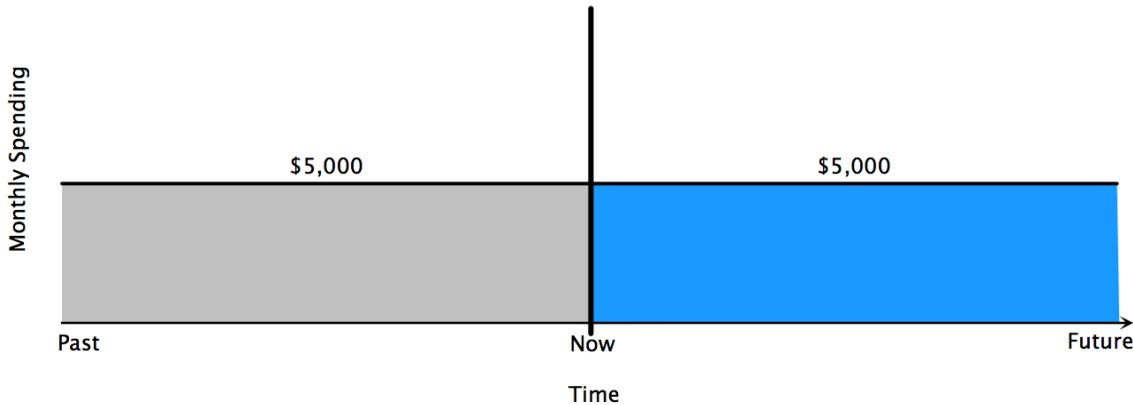


Figure 26: Instruction - Monthly Spending Graph 3

while in Universe Two your monthly spending is

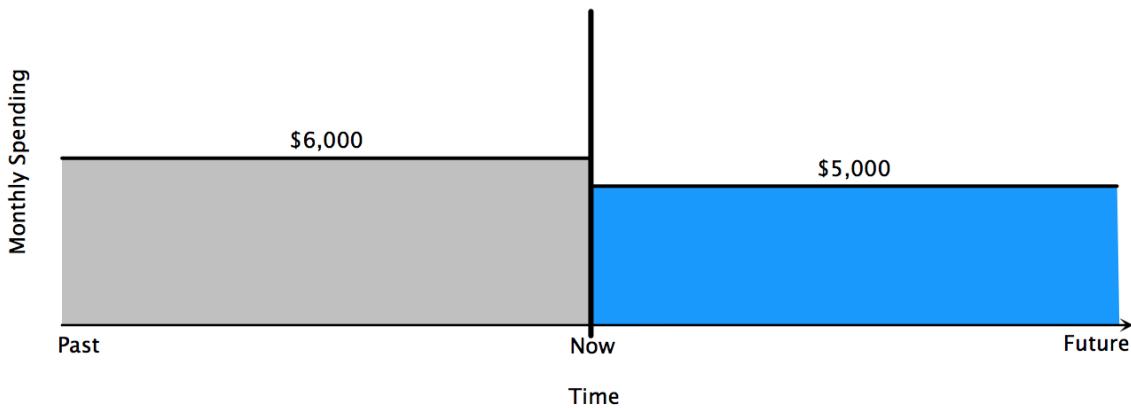


Figure 27: Instruction - Monthly Spending Graph 4

then the difference and the similarity of your monthly spending in the two universes are that

- in Universe Two you spent \$1,000 more per month in the 'past' than you did in Universe One where you spent \$5,000 per month in the 'past'; and
- in both universes, you will spend \$5,000 per month from 'now' onward.

Then, the respondents are presented with a set of practice questions testing their understanding of the above instruction.

Imagine that your monthly spending is detailed in the following monthly spending graph

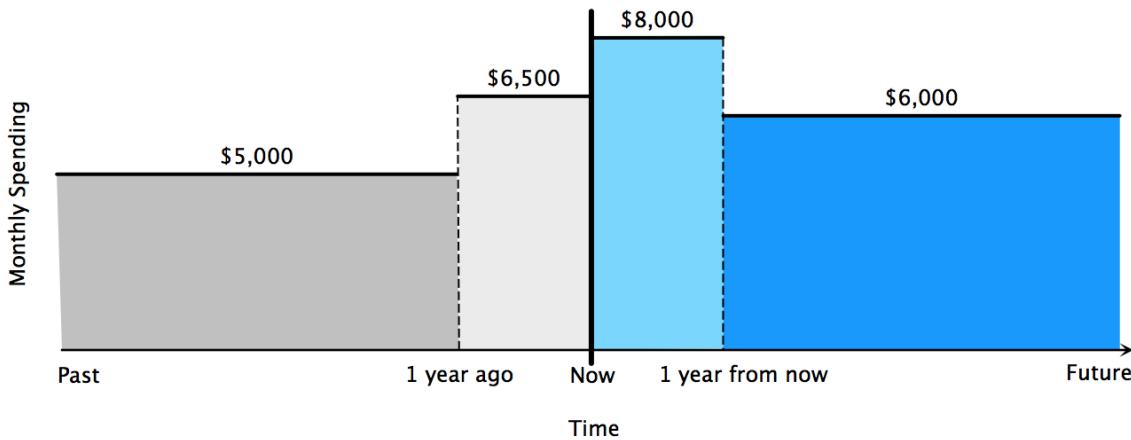


Figure 28: Instruction - Monthly Spending Graph 5

How much will you spend per month in the next year?

- \$5,000
- \$6,000
- \$6,500
- \$8,000

How much did you spend per month from 'as far back as you can remember in the past' until '1 year ago'?

- \$5,000
- \$6,000
- \$6,500
- \$8,000

- Imagine that your monthly spending in Universe One and Universe Two are detailed in the following monthly spending graphs.
- [The graphs are in the choices below. You can directly click the graph to give your answer.]

In which universe did you spend more per month in the 'past year'?

- Universe One

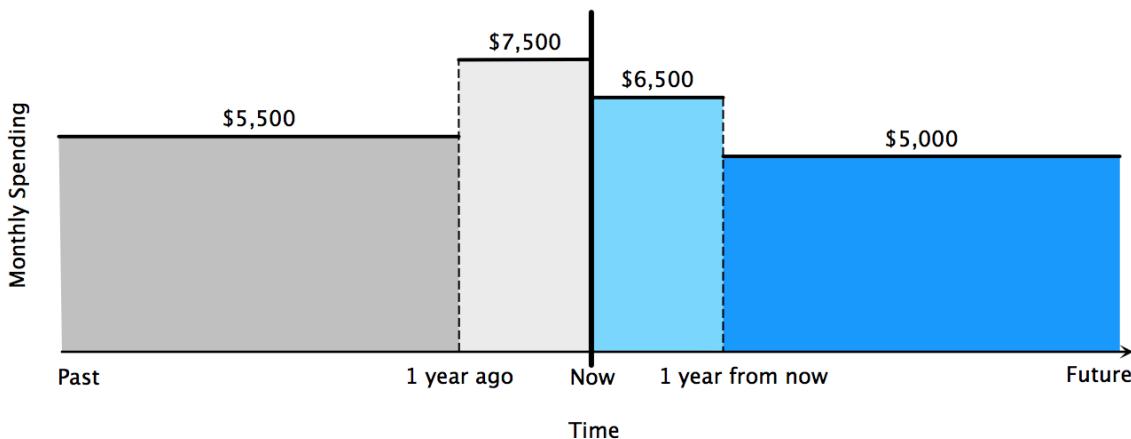


Figure 29: Instruction - Monthly Spending Graph 2

- Universe Two

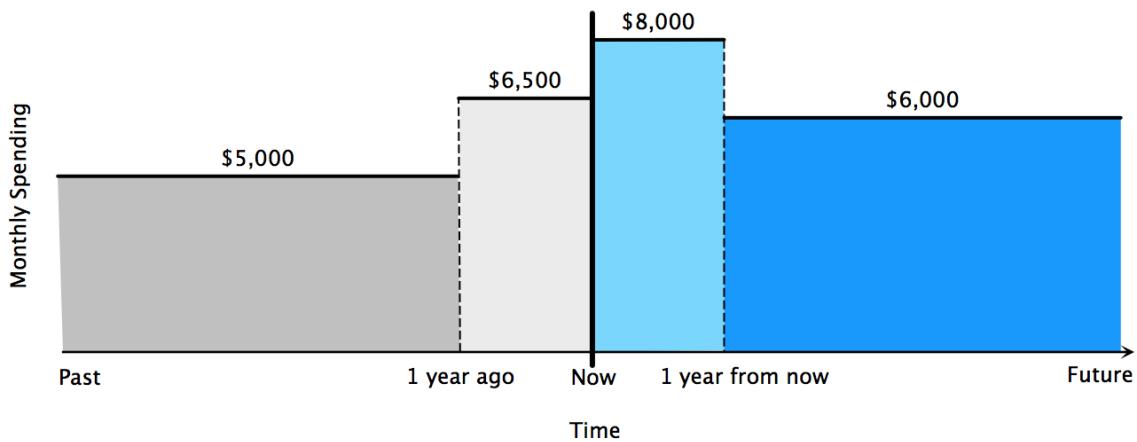


Figure 30: Instruction - Monthly Spending Graph 5

In which universe will you spend more per month from '1 year from now' onward?

- Universe One

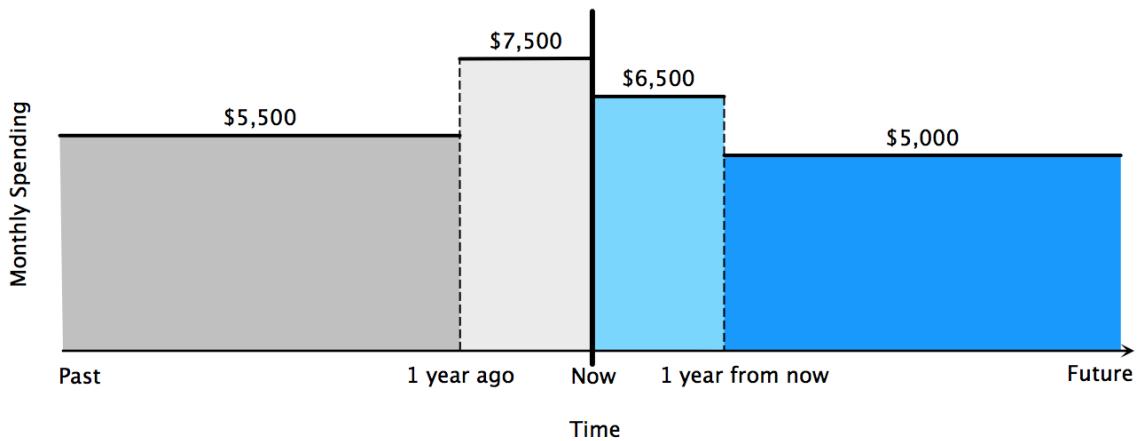


Figure 31: Instruction - Monthly Spending Graph 2

- Universe Two

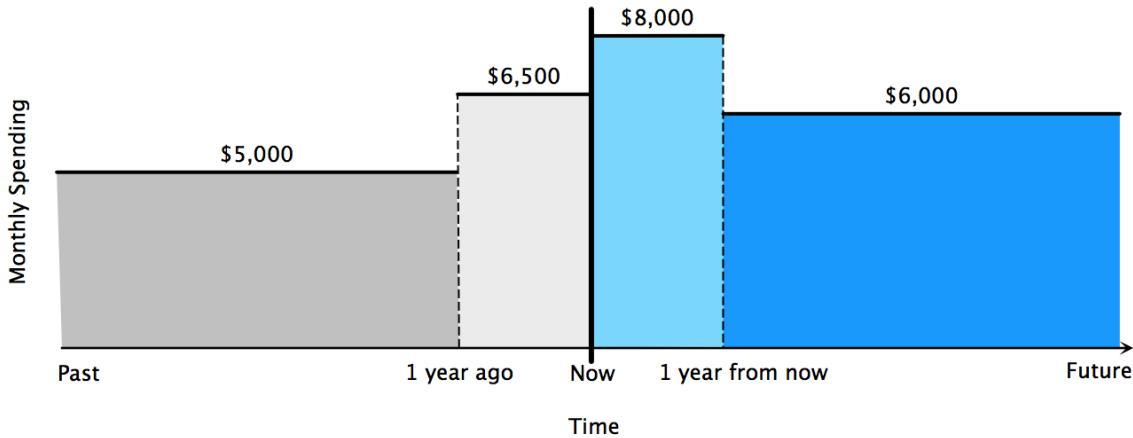


Figure 32: Instruction - Monthly Spending Graph 5

In the graphs of last question, how much more did you spend in Universe One than in Universe Two from 'as far back in the past as you can remember' until '1 year ago'?

- \$0
- \$500
- \$1,000
- \$5,500

Final Instruction

A final instruction vaccinates the respondents against the seeming repetitiveness of follow-up questions and encourages efforts with attention checks and a potential bonus reward. An opportunity to review the instructions is also presented.

We designed the survey to learn as much as possible from your answers. To increase the power of the study result, each question is normally followed by additional questions that vary slightly from previous questions. Although the survey may look repetitive, please pay careful attention and answer each question the best you can.

Implicit and explicit attention checks are integrated into this survey. Responses show signs of inattentiveness will be rejected. A small lottery (\$1) will be randomly paid as a bonus to workers who show excellence in the responses. The chance to get this lottery reward is 1 in 100. Your normal HIT payment won't be affected by this lottery.

If you would like to view these instructions again before beginning the survey, please check the following box.

View Instructions Again

F.2 Core Modules

The survey has nine core modules, each corresponding to one or two preference parameters of interest.

The flow of each core module is

1. ask the respondent to choose which universe brings her a better past experience (Figure 2);
2. ask the respondent to choose which universe brings her a better future experience (Figure 3);
3. present the respondent with a set of follow-up monthly spending graphs that varies slightly from the original set. The total number of sets of follow-up questions ranges from 1 to 2 (Figure 7).

The main element of the survey questions in the core modules is the monthly spending graphs, which are in the main text of the paper. To save space, they are not repeated here. In Table 26, I present the changes between follow-up questions and initial questions in terms of the quantities that are necessary to elicit the preference parameters in Propositions 1, 4-10, 12.

F.3 End-of-Survey Check of Understanding of Hypothetical Situation

At the end of the survey, I check the respondents' understanding of the hypothetical situation again using the following questions, which serve as implicit attention checks.

Under the hypothetical situation of this survey, if you can buy 3 bananas with one dollar in the last year, how many bananas can you buy with one dollar in the next year?

- 5
- 3
- 1
- No idea

Under the hypothetical situation of this survey, select any of the following that you own (that is, not rent):

- Residence
- Car
- Furniture

Table 26: Quantities in Monthly Spending Graphs

		If choosing U2 in initial and 1st follow-up questions (2nd follow-up question)	If choosing U2 in initial question (1st follow-up question)	Initial question	If choosing U1 in initial question (1st follow-up question)	If choosing U1 in initial and 1st follow-up questions (2nd follow-up question)
Habit Depreciation	ΔC_{U1}	400	1200	2000	2000	2000
Speed External	ΔC_{U2}	4000	4000	4000	2800	2200
Habit Mixture	ΔC	4500	1200	500	500	500
Coefficient $-u_H/u_C$	ΔC_{others}	500	500	500	1200	4500
	Δe	2000	2000	2000	2000	2000
	Δf	20	80	200	400	1000
	Δe	500	500	500	500	500
	Δf	540	600	650	700	800
	Δe	2000	1600	1000	600	100
	Δf	200	200	200	200	200
	Δe	500	1500	2200	3000	3500
	Δf	500	500	500	500	500
	ΔC_{others}^{U1}	3000	600	300	150	100
	ΔC_{others}^{U2}	3000	3000	3000	3000	3000
Time Discount	Δe	2000	2000	2000	2000	2000
Rate	Δf	3300	2500	2200	2100	2040

Notes: U1 and U2 mean Universe One and Universe Two of the monthly spending graphs, respectively. Choosing U1 in the initial question and then U2 in the 1st follow-up question or choosing U2 in the initial question and then U1 in the 1st follow-up question ends the module at the end of the 1st follow-up question (see Figure 7). All amounts are in US dollar.

- I do not own any of the above
- No idea

Under the hypothetical situation of this survey, do things you want change over time?

- Yes
- Maybe
- No

Under the hypothetical situation of this survey, do things not mentioned in the questions change?

- Yes
- Maybe
- No

Under the hypothetical situation of this survey, how much do people not mentioned in questions always spend per month?

- \$4,000
- \$5,000
- \$6,500
- \$8,000
- No idea