

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays in Peer, Time, and Risk Preferences

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy

in

Economics

by

Seung-Keun Martinez

Committee in charge:

Professor Charles Sprenger, Chair
Professor James Andreoni
Professor Michael Callen
Professor Gordon Dahl
Professor Uri Gneezy
Professor Joel Sobel

2019

Copyright

Seung-Keun Martinez, 2019

All rights reserved.

The Dissertation of Seung-Keun Martinez is approved and is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2019

TABLE OF CONTENTS

Signature Page	iii
Table of Contents	iv
List of Figures	vi
List of Tables	viii
Acknowledgements	ix
Vita	x
Abstract of the Dissertation	xi
Chapter 1 Social Comparisons in Peer Effects	1
1.1 A Theory of Social Comparisons	5
1.2 Experimental Design	8
1.2.1 The Task Choice	9
1.2.2 Self Signaling and Social Comparisons	12
1.2.3 Signal Extraction and Peer-Group Formation	13
1.3 Results	14
1.3.1 Self-Image Results	15
1.3.2 Signal Extraction Results	16
1.4 Conclusion	20
Chapter 2 Procrastination in the Field: Evidence from Tax-Filing	22
2.1 Introduction	22
2.2 Data	28
2.2.1 Tax Filing Data	28
2.2.2 Economic Stimulus Act 2008	32
2.2.3 Experimentally Elicited Time Preferences	34
2.3 Empirical Strategy	37
2.3.1 Tax Filing as Exponential Dynamic Discrete Choice	37
2.3.2 Estimating Quasi-Hyperbolic Discounting	41
2.3.3 Constructing Counterfactuals	43
2.4 Results	45
2.4.1 Exponential (δ) Discounting, Quasi-Hyperbolic (β - δ) Discounting, and Excess Mass	45
2.4.2 Response to Changing Incentives: 2008 Stimulus Payments	53
2.4.3 Experimental Present Bias and Tax Filing Behavior	56
2.5 Discussion and Conclusion	59
Chapter 3 A Test for Risk-Averse Expected Utility	62

3.1	Introduction	62
3.2	The Model	66
3.3	Conclusion	78
Appendix A Social Comparisons in Peer Effects		80
A.1	Results	80
A.2	Experiment Instructions	85
A.3	News Article about the A.D.R.P.	95
Appendix B Procrastination in the Field		97
B.1	Additional Tables and Figures	97
B.2	Tax Filing as Dynamic Discrete Choice	107
B.2.1	Type-1 Extreme Value Errors	108
B.2.2	Likelihood Formulation	112
B.3	Refund Arrival Delay and Identification	114
B.3.1	Starting and Limiting Values	114
B.3.2	Single Coincidence	115
B.3.3	Level Curve Simulations	117
B.4	Naive $\beta - \delta$ derivation	120
B.5	IRS Announcement of H.R. 5140	123
B.6	Excerpts from Technical Explanation of H.R. 5140	124
B.6.1	Explanation of Provision	124
B.6.2	In general	124
B.6.3	Basic credit	124
B.6.4	Qualifying child credit	125
B.6.5	Limitation based on gross income	125
B.7	Instructions	127
References		129

LIST OF FIGURES

Figure 1.1:	Concavity Induced Peer Effects	6
Figure 1.2:	Social Information Anticipation	8
Figure 1.3:	Example Task	10
Figure 1.4:	Decision Screen	11
Figure 1.5:	Mean Task Choice by Donation	15
Figure 1.6:	Average Tasks Choice by Signal, Treatment	18
Figure 1.7:	Distribution of Beliefs by Signal, Treatment	20
Figure 2.1:	Filing Times and Refund Values	46
Figure 2.2:	Predicted and Actual Filing Behavior	49
Figure 2.3:	Filing Dates, Stimulus Payments, and Out-of-Sample Predictions	55
Figure 2.4:	Predicted and Actual Filing Probabilities by Present Bias	58
Figure 3.1:	η matrix	76
Figure A.1:	Distribution of Task Choice by Treatment	80
Figure B.1:	Histogram of 2008 Stimulus Act Payments	97
Figure B.2:	Predicted Filing Times Using 1-Period-Ahead Observed Probabilities	101
Figure B.3:	Filing Times and Refund Values (Paper Check Only)	102
Figure B.4:	Filing Times and Refund Values (Direct Deposit Only)	102
Figure B.5:	Filing Times and Refund Values (δ Heterogeneous by Race)	103
Figure B.6:	Filing Times and Refund Values (Present Biased, Paper Check Only)	103
Figure B.7:	Filing Times and Refund Values (Present Biased, Direct Deposit Only)	104
Figure B.8:	Filing Times and Refund Values 2008 (Exponential)	104
Figure B.9:	Filing Times and Refund Values 2008 (Present Biased)	105
Figure B.10:	Predicted and Actual Filing Behavior by Present Bias (Un-smoothed)	106

Figure B.11:	Conditional Probability Level Curves assuming $\delta = 0.535, c = 10$	119
Figure B.12:	Conditional Probability Level Curves assuming $\delta = 0.535, c = 10$	119
Figure B.13:	Conditional Probability Level Curves assuming $\delta = 0.535, c = 10$	120

LIST OF TABLES

Table 1.1:	Effect of Treatment on Tasks Allocation	17
Table 1.2:	Task Choice by Received Signal	19
Table 2.1:	Summary Statistics	33
Table 2.2:	Aggregate Parameter Estimates 2005 — 2007	47
Table 2.3:	Excess Mass Results	51
Table 2.4:	Predicted and Actual Difference-in-Difference Effect of Stimulus Payments	56
Table 2.5:	Present Bias and Deviations from Exponential Prediction	59
Table A.1:	Distribution of Task Choice by Treatment	81
Table A.2:	Task Choice by Belief, Signal, Treatment	82
Table A.3:	Effect of Treatment on Likelihood of Choosing At Least Twenty Tasks	82
Table A.4:	Quantile Regressions	83
Table A.5:	Quantile Regressions	83
Table A.6:	Small Sample t-test: Effect of High Signal on Task Choice	84
Table A.7:	Effect of Treatment on Belief of Received Signal	84
Table B.1:	Summary statistics	98
Table B.2:	Aggregate Parameter Estimates 2007	98
Table B.3:	Aggregate Parameter Estimates 2005 — 2007, $\mu = 0$	99
Table B.4:	Aggregate Parameter Estimates 2005 — 2007, $k = 0$	99
Table B.5:	Aggregate Parameter Estimates 2005 — 2007, Paper Check	100
Table B.6:	Aggregate Parameter Estimates 2005 — 2007, Direct Deposit	100
Table B.7:	Aggregate Parameter Estimates 2005 — 2007, Exclude Last 7 Days	101
Table B.8:	Difference-in-Difference Effect of Stimulus Payments (Placebo)	105

ACKNOWLEDGEMENTS

I would like to first acknowledge my dissertation advisor, Charlie Sprenger. He is an deeply generous advisor and was always willing to share his encyclopedic knowledge and sharp intuition for behavioral economics with me and his other students. I would also like to thank several other advisors and collaborators who have helped me become the researcher I am today. John Morgan was crucial to my early development as an economist and helped direct me to enroll in this PhD program. Gordon Dahl provided instrumental guidance on my research agenda on peer effects. Jason Bigenho is an ever-dependable friend and collaborator. Others who deserve specific mention include Jim Andreoni, Ned Augenblick, Mike Callen, Chris Chambers, Carlos Cueva, Ce Liu, Uri Gneezy, Stephan Meier, and Joel Sobel.

Chapter 1, in full, is currently being prepared for submission for publication of the material. Martinez, Seung-Keun; Bigenho, Jason. “Social Comparisons in Peer Effects.” The dissertation author was one of two primary investigators and authors of this material.

Chapter 2, in full, is currently being prepared for submission for publication of the material. Martinez, Seung-Keun; Charles, Sprenger; Stephan, Meier. “Procrastination in the Field: Evidence from Tax-Filing” The dissertation author was one of three primary investigators and authors of this material.

Chapter 3, in full, has been published. Martinez, Seung-Keun; Chambers, Christopher; Liu, Ce. “A Test for Risk-Averse Expected Utility.” The dissertation author was one of three primary investigators and authors of this material.

VITA

- 2013 Bachelor of Arts, Mathematics-Economics, University of California San Diego
- 2014 Master of Arts, Economics, University of California San Diego
- 2019 Doctor of Philosophy, Economics, University of California San Diego
- 2019– Assistant Professor of Economics, University of Nottingham

ABSTRACT OF THE DISSERTATION

Essays in Peer, Time, and Risk Preferences

by

Seung-Keun Martinez

Doctor of Philosophy in Economics

University of California San Diego, 2019

Professor Charles Sprenger, Chair

This dissertation is composed of three papers on distinct topics, each studying a different aspect of decision making. In chapter 1, we study why, even without material incentives for coordination or learning, social interactions influence individual decision making. Identifying why conformity arises absent explicit incentives faces the challenge that any rationalizing theory must rely on unobservable preferences or beliefs. Therefore, empirical distinction requires theories that make predictions beyond the basic dynamic of conformity. To that end, we propose and test a theory of self signaling in peer effects. The model generates (partial) conformity as a response to how others' choices inform one's own self-image. The mechanism of self-signaling for peer effects delivers unique, falsifiable predictions that we test experimentally. The theory

predicts that the anticipation of learning others' choices and the garbling of information on others' choices will both deeply influence behavior. In two real-effort lab experiments we not only find treatment effects closely in line with the model's unique predictions, but also document the importance of self image in social comparisons.

In chapter 2, we attempt to identify present-biased procrastination in tax filing behavior. Our exercise uses dynamic discrete choice techniques to develop a counterfactual benchmark for filing behavior under the assumption of exponential discounting. Deviations between this counterfactual benchmark and actual behavior provide potential 'missing-mass' evidence of present bias. In a sample of around 22,000 low-income tax filers we demonstrate substantial deviations between exponentially-predicted and realized behavior, particularly as the tax deadline approaches. Present-biased preferences not only provide qualitatively better in-sample fit than exponential discounting, but also have improved out-of-sample predictive power for responsiveness of filing times to the 2008 Economic Stimulus Act recovery payments. Additional experimental data from around 1100 individuals demonstrates a link between experimentally measured present bias and deviations from exponential discounting in tax filing behavior.

In chapter 3, we provide a universal condition for rationalizability by risk-averse expected utility preference in a demand-based framework with multiple commodities. Our test can be viewed as a natural counterpart of a classical test of expected utility, due to Fishburn (1975), in a demand setting.

Chapter 1

Social Comparisons in Peer Effects

The importance of social information in individual decision making is well documented. Not only do individuals learn from their peers when making decisions over new and unfamiliar opportunities (Foster and Rosenzweig, 1995; Duflo and Saez, 2002; Bursztyn et al., 2014; Dahl, Løken and Mogstad, 2014), but they also converge to behavioral conformity in the work place (Mas and Moretti, 2009; Bandiera, Barankay and Rasul, 2010) and in the classroom (Hoxby, 2000; Sacerdote, 2001; Zimmerman, 2003; Bursztyn, Egorov and Jensen, 2017). While material incentives for coordination or reliance on social information for uncertainty resolution can explain some instances of social conformity (Katz and Shapiro, 1986; Banerjee, 1992; Foster and Rosenzweig, 1995), previous research demonstrates the importance of social influence that is not predicted by neo-classical theories (Frey and Meier, 2004; Falk and Ichino, 2006; Goette, Huffman and Meier, 2006; Alpizar, Carlsson and Johansson-Stenman, 2008; Mas and Moretti, 2009; Shang and Croson, 2009; Chen et al., 2010).

Distinguishing between plausible mechanisms for conformity in the absence of explicit incentives faces a natural challenge. If incentives are unobservable, any rationalizing theory will necessarily rely on unobservable preferences or beliefs. For example, a common explanation for such peer effects is adherence to social norms. These theories are typically driven by the stigma of deviating from a socially prescribed action (Akerlof, 1980; Jones, 1984; Akerlof and

Kranton, 2000) or by individuals pooling at a common action to signal an optimal social type (Bernheim, 1994).¹ While we may be able to observe individuals make similar choices, it is difficult to observe whether they do so specifically because it is expected of them. Therefore, theories of peer effects are empirically distinguishable only if they differ in their predictions beyond the basic dynamic of conformity begetting further conformity. In this project we propose and experimentally test a theory of (partial) social conformity in the absence of a socially optimal action. Our model produces a novel prediction on how decision makers will alter their choices in anticipation of learning the choices of others. Further, we introduce a theoretical foundation and experimental test for self image as a mechanism underlying peer effects.

Our theory posits that individuals have an intrinsic desire to judge and evaluate themselves. That is, each person would like to perceive himself positively—for example, as diligent, intelligent, and charitable. However, individuals often lack a direct or objective means of self evaluation. Instead, they rely on their history of actions as a noisy signal of their attributes. This theory requires no predetermined socially optimal action. Rather, self-image is increasing in the performance of a costly action. Under these assumptions an individual faces intrinsic incentives to manipulate his personal image through his actions. Further, signal extraction is improved when he observes others performing similar tasks—i.e. social comparisons allow him to better understand the image implications of his own choices.

We find that if self image is increasing in one's own performance and decreasing in peer performance, then diminishing marginal utility over self image will produce positive peer effects. Intuitively, concave image utility implies the shame of learning you are the lowest social type is greater than the pleasure of learning you are the highest social type. As such, individuals are content following the majority. More concretely, consider a group of employees working on an unfamiliar task under a fixed-wage spot contract that offers no incentives for collusion on effort.² Why would an employee condition his own output on the observed output of a peer? Suppose

¹Bernheim and Exley (2015) explore an alternative explanation for social conformity that establishes preference mechanisms that drive instances of social conformity. Our project explores a belief mechanism.

²Previous experiments have documented peer effects under such conditions (Falk and Ichino, 2006).

that each employee wishes to perceive himself as hard-working but is unsure how to judge his performance. If each employee draws comparisons with peer output to better understand whether his output suggests diligence or laziness, then our theory establishes that a group of employees will conform in output when each individual abhors being the laziest group member more than he enjoys being the most diligent.

Modeling social comparisons as self signaling offers novel predictions beyond the basic dynamic of conformity. In particular, our model predicts individuals will make more costly but more image-enhancing choices in anticipation of social information. In the context of our previous example, suppose that no one observes his coworkers' donations before he chooses his own donation. In our model, a donor's potential marginal utility loss from learning all his coworkers gave more than him is greater than the potential utility gain of learning he gave more than his coworkers. Therefore, each donor hedges against bad news and gives more if he knows we will eventually learn his coworkers' donations.

The experiments in this project corroborate our theory's prediction on self-image-produced conformity. The first experiment uses the model's social information anticipation prediction to test the relevance of self image in peer effects. Specifically, we test whether individuals are willing to transcribe more blurry images of text in exchange for charitable donations if they know they will learn the distribution of previous participant choices than when they know they will remain uninformed. Importantly, all participants commit to how many of these transcriptions they are willing to complete prior to learning any information on how their decisions compare to the decisions of others. Both those who will and will not be shown the distribution of others' decisions must predict where their decision lies in the distribution of all previous choices. Subjects are also informed of exactly what the experimenter observes—the full anonymous distribution of all participant choices—in both treatment and baseline. Therefore, we ensure that any treatment effect is driven entirely by each participant's anticipation of learning how his choice compares to the decisions of others and not by what anyone else learns.³ In accor-

³To that end, subjects did not learn the choices of others in their sessions, and subjects in both treatment

dance with our theory's predictions, we find that participants are willing to do more tasks when they anticipate learning how their decisions compare to the decisions of previous participants

The second experiment tests whether a noisier signal mitigates peer effects by statistically garbling the information peer behavior provides. Documenting that decision makers are less responsive to less precise signals is an important result of this project. Our theory relies on individuals reacting to their expectations of peer behavior to better understand the image implications of their own choices. For this theory to be plausible, not only must the basic dynamics of conformity hold, but participants must be sensitive to the receipt of a statistically refined signal on peer choices and unresponsive to an uninformative signal on peer choices.

The economic scope of peer influence is vast. Previous work has documented that people's uptake of retirement savings programs, charitable giving, effort at work, managerial strategies, decisions to invest in financial instruments, and participation in paternal leave are causally related to the observed decisions of their peers (Duflo and Saez, 2002; Frey and Meier, 2004; Mas and Moretti, 2009; Bandiera, Barankay and Rasul, 2010; Shue, 2013; Bursztyn et al., 2014; Dahl, Løken and Mogstad, 2014), and that students' academic achievements covary with those whom they share a classroom or dormitory (Hoxby, 2000; Sacerdote, 2001; Zimmerman, 2003). A careful investigation of how self-signaling and social comparisons may drive social conformity could not only contribute to our understanding of the existing empirical and theoretical literature, but also provide researchers and policy makers with greater predictive power over when peer effects are likely to exist in unexplored environments.

Section 1.1 presents a simple model of social comparisons and explains the intuition behind our results. Section 1.2 details our experimental design, and section 1.3 discusses the results of our experiment. Finally, section 1.4 concludes and discusses possible future work on this topic.

and baseline are informed that the anonymous distribution of previous participant choices may be used in future experiment sessions.

1.1 A Theory of Social Comparisons

In this section, we present a simple model of self image and social comparisons. This model serves two primary purposes. First, we use this model to deliver sufficient conditions for conformity. Second, the model produces a testable prediction on how individuals will make self-image-relevant choices in anticipation of learning the choices of others.

Our first results states that if an individual experiences positive image returns in the performance of a costly action and and negative image returns in the observation of higher peer performance, then diminishing marginal utility over self image will produce convergence in group behavior. To see this, suppose that individual 1 takes a costly action a_1 and observes the costly action of individual 2, a_2 . Define an image function $I(a_1, a_2)$ for individual 1 such that I is increasing in a_1 and decreasing in a_2 : $\frac{\partial I}{\partial a_1} > 0 > \frac{\partial I}{\partial a_2}$. Lastly, let person 1 experience utility U over image I . For example, a_1 and a_2 may represent charitable donations. $I(a_1, a_2)$ is agent 1's perception of his own altruism when he compares his gift to the gift of agent 2. $U(I)$ is then the utility agent 1 experiences over how altruistic he believes himself to be.

Actions a_1 and a_2 incur costs $C(a_1)$ and $C(a_2)$. Where $C' > 0$ and $C'' \geq 0$. Naturally, agent 1 will choose a_1 by equating marginal image utility to the marginal cost of activity. Under these assumptions agent 1 will choose higher a_1 in response to higher a_2 if image utility U is increasing and concave.

The geometric intuition behind our first result is shown in shown in figure 1.1. Let $a_2 + h > a_2$, and, for simplicity, assume that the image function I is linear. Under these conditions we have that $U'(I(a_1, a_2 + h)) \frac{\partial I}{\partial a_1} > U'(I(a_1, a_2)) \frac{\partial I}{\partial a_1}$ for all a_1 due to the strict concavity of U . Therefore, since the cost function $C(a_1)$ is not dependent on a_2 , we have that $\text{argmax}_{a_1} U(I(a_1, a_2 + h)) - C(a_1) > \text{argmax}_{a_1} U(I(a_1, a_2)) - C(a_1)$. In other words, person 1 chooses more costly a_1 in response to a more costly choice of a_2 because the marginal image utility of activity at $I(a_1, a_2 + h)$ is higher than the marginal image utility at $I(a_1, a_2)$.

Importantly, our model also predicts that self signaling will induce decision makers to

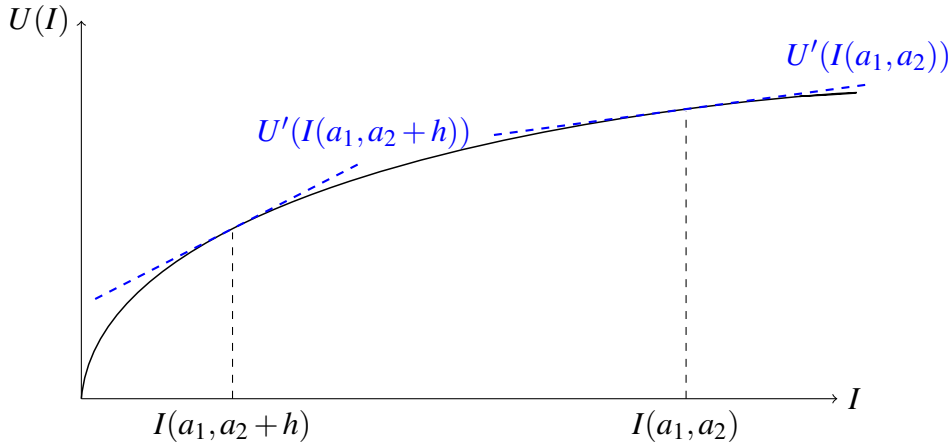


Figure 1.1: Concavity Induced Peer Effects

choose more costly actions in anticipation of learning how their choices compare to the decisions of others. This prediction crucially relies on the assumption that utility over self image is strictly increasing and diminishingly concave – i.e. $U' > 0$, $U'' < 0$ and $U''' > 0$.

The intuition behind this result is similar to that of precautionary savings under prudent expected utility risk preferences. Recall that under expected utility preferences, the prudent decision maker saves more money today in response to an increase in future income uncertainty. That is, prudent expected utility preferences predict that an individual will choose to save against a possible future negative income shock. Our model of self image makes a similar prediction on how individuals will make choices in anticipation of learning the choices of others. In our model, the revelation of future social information causes a mean-preserving spread of one's future self image. The decision maker will then work harder today to bolster his self image to hedge against a possible negative image shock tomorrow.

Formally, suppose that agent 1 makes his choice a_1 knowing that agent 2 has already made her choice. Let $a_i \in [0, A]$, and suppose that agent 1 holds a prior over agent 2's, $a_2 \sim F_2$. Where we assume F_2 differentiable. Then we will show, that agent 1 will make a more costly and image-enhancing choice in anticipation of learning whether $a_1 > a_2$ or $a_1 < a_2$.

The mechanism behind this result is shown in figure 1.2. Let $I(a_1, E[a_2])$ be agent 1's

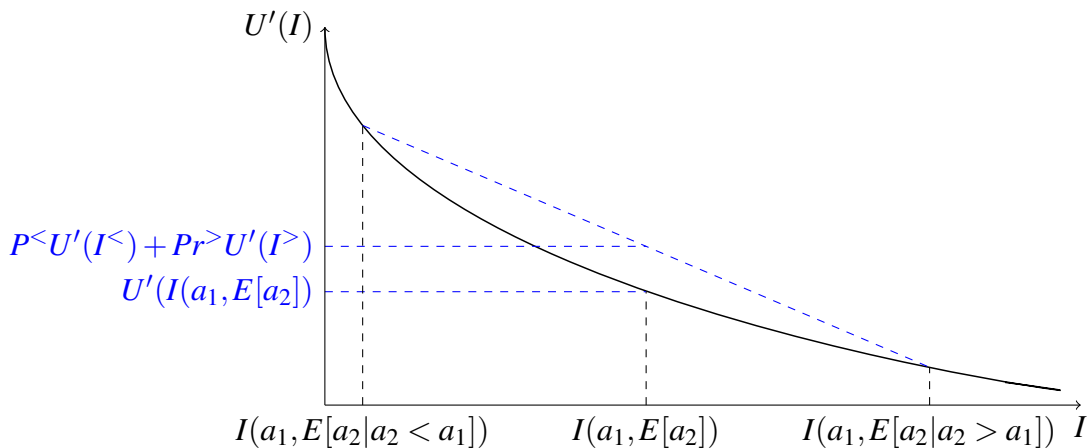
image when he does not learn agent 2's choice. Let $I(a_1, E[a_2|a_1 < a_2])$ be agent 1's image if he learns $a_1 < a_2$. Similarly, let $I(a_1, E[a_2|a_1 > a_2])$ be agent 1's image if he learns that agent 2 chose a less costly action. Then the optimization problem in which agent 1 anticipates learning whether $a_1 \leq a_2$, but only after he makes his decision, is $\max_{a_1} Pr(a_1 < a_2|a_1)U(I(a_1, E[a_2|a_1 < a_2])) + (1 - Pr(a_1 < a_2|a_1))U(I(a_1, E[a_2|a_1 > a_2])) - C(a_1)$. Notice that promise of learning whether $a_1 \leq a_2$ introduces a mean preserving spread to agent 1's image.⁴ Since we assume that $U''' > 0$, by Jensen's inequality we have that $Pr(a_1 < a_2|a_1)U'(I(a_1, E[a_2|a_1 < a_2])) + (1 - Pr(a_1 < a_2|a_1))U'(I(a_1, E[a_2|a_1 > a_2])) > U'(I(a_1, E[a_2|a_1 > a_2]))$ for all a_1 .

As such,

$$\begin{aligned} & \operatorname{argmax}_{a_1} Pr(a_1 < a_2|a_1)U(I(a_1, E[a_2|a_1 < a_2])) + \\ & (1 - Pr(a_1 < a_2|a_1))U(I(a_1, E[a_2|a_1 > a_2])) - C(a_1) > \operatorname{argmax}_{a_1} U(I(a_1, E[a_2|a_1 > a_2])) - C(a_1) \end{aligned}$$

That is, the decision maker will choose a more costly task when he knows that he will learn how his choice compares to the choices of others than when he is to remain ignorant to others choices. This occurs because convexity in marginal image utility implies that the potential opportunity cost of choosing too low of an action when $a_1 < a_2$ outweighs the opportunity cost of choosing too high of an action when $a_1 > a_2$.

⁴ $Pr(a_1 < a_2|a_1) \int_0^{a_1} \frac{A_2 f_2(A_2)}{Pr(a_1 < a_2|a_1)} dA_2 + (1 - Pr(a_1 < a_2|a_1)) \int_{a_1}^A \frac{A_2 f_2(A_2)}{(1 - Pr(a_1 < a_2|a_1))} dA_2 = \int_0^A A_2 f_2(A_2) dA_2$ by definition, and we assumed that I is linear.



Where $I^< = I(a_1, E[a_2|a_2 < a_1])$, $I^> = I(a_1, E[a_2|a_2 > a_1])$, $P^> = Pr(a_1 > a_2)$, and $P^< = Pr(a_1 < a_2)$

Figure 1.2: Social Information Anticipation

1.2 Experimental Design

We conducted two experiments to directly test the predictions of our theory. The first experiment tested the relevance of self image in social comparisons, and the second experiment tested the relevance of signal extraction. In both experiments participants chose the maximum number of real-effort tasks they were willing to complete in exchange for a donation to charity. All experiment sessions were conducted at the University of California, San Diego Department of Economics. Subjects completed the experiment on individual computer terminals. Privacy screens were installed on each computer and barriers were placed between every subject. All participants were current undergraduates at UC San Diego. Each subject was paid 15 dollars for his or her participation, and experiment sessions took approximately 50 minutes. Treatments were varied across sessions for a between-subjects design. Recruitment for each session was done by emailing a random sample of UCSD undergraduates. The first experiment contained nine to fourteen subjects per session, and the second experiment contained eight to twelve subjects per session.

Section 1.2.1 explains the real-effort task choice. Section 1.2.2 details the self-image experiment, and section 1.2.3 describes the test for signal extraction.

1.2.1 The Task Choice

In both experiments, participants were asked to select the maximum number of real-effort tasks they were willing to complete in exchange for donations to the Afghan Dental Relief Project.⁵ The ADRP is a charitable organization that provides free dental services and dental health education to the poorest families and individuals in Kabul, Afghanistan. 100% of the donations that were generated from the experiment were used to purchase dental supplies for ADRP's free dental clinic. We collaborated with the ADRP because it is a deserving and relatively unknown charity. The founder, Dr. James Rolfe, agreed to monitor gifts made to the charity during the course of our experiment. Dr. Rolfe reports that no gifts were made to the ADRP that could have been given by a UCSD undergraduate.⁶

The real-effort task was to transcribe captchas.⁷ Captchas are distorted images of a sequence of letters commonly used by website developers to distinguish between human users and bots. The task was deliberately made to be more frustrating and tedious than typical website captchas. We wanted subjects to experience increasing marginal costs in their task performances. Our captchas consisted of capital and lowercase letters, numbers, and special characters. Only correctly transcribed captchas were considered completed tasks. Subjects were given three chances per randomly assigned captcha, and they could not skip an assigned captcha except by deliberately entering three incorrect responses.⁸ The typical participant was able to correctly transcribe 1-out-of-every-5 captchas. An example of the task is show in Figure 1.3.⁹

In the self-image experiment subjects selected the maximum number of tasks they were

⁵We elicited willingness to do real-effort tasks in exchange for donations to charity, as opposed to eliciting willingness to give money, so that we could measure how willingness to give one's time and effort changes with the size of the donation.

⁶Specifically, all donations made to the ADRP during the course of our experiment came from us, the experimenters, or from members of Dr. Rolfe's local Santa Barbara community.

⁷Captchas were generated using the Python module `Claptcha`, available here: <https://github.com/kuszaj/claptcha>. All use of this module is in accordance with the license outlined at the previous link. The font used for captchas was "Mom's Typewriter," an open-source font available here: <https://www.dafont.com/moms-typewriter.font>.

⁸All captchas were randomly assigned out of a bank of 1000 captcha images.

⁹The correct transcription of the shown captcha is "4LBMwW?E"

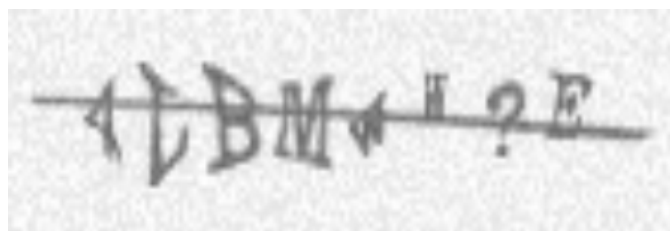


Figure 1.3: Example Task

willing to complete in exchange for donations of \$2, \$5, \$10, \$15, and \$20. Each subject chose between 0 and 50 tasks for each possible donation amount. We ensured the incentive compatibility of this decision by randomly assigning each subject a donation and task amount after they had completed making their decisions. If the assigned number of tasks was equal to or fewer than the maximum number they were willing to do for their assigned donation, they completed their assigned number of tasks and we gave the assigned donation to the ADRP. For example, suppose that for a donation of \$15 a participant chose to do at most 30 tasks. If he were randomly assigned \$15 and 20 tasks, then he completed 20 tasks and \$15 was given to the Afghan Dental Relief Project. On the other hand, if he was randomly assigned \$15 and 40 tasks, then he did 0 tasks and no donation was given to the ADRP. After completing their assigned number of tasks, subjects were all paid the same participation fee as they exited the lab. Subjects were offered the same choice in the signal-extraction experiment. However, only a donation of \$20 was possible. Therefore, subjects in the signal-extraction experiment only made one decision while subjects in the self-image experiment made five.¹⁰

We wanted all participants to make an informed choice on how many tasks they were willing to complete in exchange for a donation to charity. Therefore, all subjects read a news article about the ADRP and attempted to transcribe five sample captchas prior to their decisions. Additionally, subjects received instruction and comprehension testing about the incentive

¹⁰Each subject was assigned a donation amount by drawing a ball out of a jar. Each ball had a dollar amount and three-digit code written on it. The codes ensured that no subject could choose to work for a donation amount that they were not assigned. A random number generator then assigned each subject a number of tasks (between 0 and 50). Both randomization devices were fair – every donation and task combination had an equal probability of being assigned.

compatible task choice. We wanted to prevent contemporaneous peer effects in our experiment. Therefore, individuals never observed the decisions of others in their own session, nor did they observe the randomly assigned donations or task amounts of other participants in their session. Lastly, we wanted to measure how many captchas participants wanted to complete, not how many they thought they could complete within a specific time interval. Therefore, subjects were given as much time as they needed to complete their tasks.

Task Choices

Please indicate the maximum number of tasks you are willing to do for each donation amount. Please also guess what percent of other experiment participants were willing to do more tasks than you.

Donation Amount	Your Choices
\$2	<input type="range" value="0"/> 0 What percent of previous participants do you think were willing to do more tasks than you for a \$2 donation? <input type="radio"/> 95% <input type="radio"/> 75% <input type="radio"/> 50% <input type="radio"/> 25% <input type="radio"/> 0%
\$5	<input type="range" value="0"/> 0 What percent of previous participants do you think were willing to do more tasks than you for a \$5 donation? <input type="radio"/> 95% <input type="radio"/> 75% <input type="radio"/> 50% <input type="radio"/> 25% <input type="radio"/> 0%
\$10	<input type="range" value="0"/> 0 What percent of previous participants do you think were willing to do more tasks than you for a \$10 donation? <input type="radio"/> 95% <input type="radio"/> 75% <input type="radio"/> 50% <input type="radio"/> 25% <input type="radio"/> 0%
\$15	<input type="range" value="0"/> 0 What percent of previous participants do you think were willing to do more tasks than you for a \$15 donation? <input type="radio"/> 95% <input type="radio"/> 75% <input type="radio"/> 50% <input type="radio"/> 25% <input type="radio"/> 0%
\$20	<input type="range" value="0"/> 0 What percent of previous participants do you think were willing to do more tasks than you for a \$20 donation? <input type="radio"/> 95% <input type="radio"/> 75% <input type="radio"/> 50% <input type="radio"/> 25% <input type="radio"/> 0%

[Next](#)

Notes: This is the decision screen for participants in the self-image experiment.

Figure 1.4: Decision Screen

1.2.2 Self Signaling and Social Comparisons

The first experiment tested our theory’s prediction that individuals will choose more costly actions in anticipation of social information. In the context of our experiment, our theory predicted that individuals would be willing to complete more captchas in anticipation of learning how their choices compared to the choices of previous participants. We used this prediction to identify the role of self image in social comparisons.

In both treatment and baseline, participants chose the maximum number of captchas they were willing to complete in exchange for donations of \$2, \$5, \$10, \$15, and \$20—as described in section 1.2.1. In conjunction with their choices, subjects were also asked to predict the percent of all previous participants that were willing to complete more tasks than them for each of their decisions. This decision is depicted in Figure 1.4. As Figure 1.4 shows, subjects selected the maximum number of tasks they were willing to complete and guessed whether 95%, 75%, 50%, 25%, or 0% of all previous participants were willing to do more tasks than them. For example, suppose that for a donation of \$10 dollars an individual chose to do a maximum of 12 tasks. Further suppose that this person thought that at least 50%, but fewer than 75%, of all previous participants were willing to do more than 12 tasks for a \$10 donation. This person was then instructed to indicate that he believes that at least 50% of all previous experiment participants were willing to do more tasks than him for a \$10 donation.¹¹

In treatment sessions, participants were given sealed envelopes that they opened *after* submitting their decisions and distributional guesses. These envelopes contained the true 5th, 25th, 50th, 75th, and 100th percentiles for task choice by donation amount. Treatment participants then submitted the correct answers for what percent of previous participants were willing to do more tasks than them for each possible donation.

All choices were recorded anonymously. No subject choice, assigned donation, or assigned task amount was revealed to others in their sessions. All subjects were aware that

¹¹These guesses were *not* incentivized to preclude subjects from tailoring their choices to their guesses so as to receive higher payment—for example, picking corners on their choices for more accurate guesses.

exactly 50% of all experiments session would learn the distribution of previous participant choices. Further, while only treatment subjects ultimately learned the percentiles of previous participant choices, we deliberately asked all participants to consider how their choices compare to the choices of all previous participants. By doing so, we made subjects in both baseline and treatment consider how their choices may compare to the choices of others. Further, we ensured all experiment subjects were aware that the experimenter will observe the full anonymous distribution of participant choices. Since the experimenter’s information set is identical in treatment and baseline, any difference in subject behavior is solely due to subjects anticipating what they themselves will learn.¹²

1.2.3 Signal Extraction and Peer-Group Formation

All of the predicted dynamics of our proposed theory rely on signal extraction. Therefore, peer effects are predicted to arise only if group behavior provides a sufficiently refined signal. In the context of our experiment, our theory states that decision makers may refer to others’ choices to better understand whether their own choices reveal selfishness or task difficulty. As such, we use statistically garbled and refined social information to directly test the relevance of statistical reasoning in social comparisons.

Recall that, participants in this experiment chose the maximum number of captchas they were willing to complete in exchange for a \$20 donation to the ADRP. Prior to their decision, each person was shown both of the following statements on the choices of all previous participants.

1. *More than 50%* of all previous participants were willing to complete at least 20 tasks for a

¹²We were also concerned that if only treatment subjects were aware that their anonymous choices may be revealed in future sessions, then social image could still be a confounding factor. Similarly, there is a second order equilibrium concern that if all subjects in the baseline thought all previous participants were also in the baseline—and all treatment participants thought all previous participants were in the treatment—then subjects in the treatment may have had higher beliefs about what previous participants chose to do. *However, we preclude these possible concerns by explaining to subjects in both treatment and baseline that 50% of all experiment sessions receive social information and 50% do not.* At no point do we reveal how many sessions have taken place or how many remain. Therefore, all subjects were uniformly in the dark about the probability that their choices will inform future revelations of social information.

donation of \$20.

2. *Less than 25%* of all previous participants were willing to complete at least 20 tasks for a donation of \$20.

Prior to making their own choices each subject also received a signal on which of these statements was true. They received this signal by drawing one of sixteen available envelopes at random. In the baseline, 50% (8-of-16) envelopes contained the true statement, while in treatment 94% (15-of-16) envelopes contained the true statement.¹³ Clearly, previous participant choices do not depend on the drawn statement of any current participant. Therefore, all that changes between treatment and baseline is the underlying probability of drawing a true or false signal. After reading their signals and choosing how many tasks they were willing to complete, subjects also indicated which signal they believed to be true.¹⁴

Signal extraction has a specific hypothesis in this context. Baseline subjects should understand that variation in the obtained signal is pure noise¹⁵, and their task choices and beliefs should be independent of the obtained signal. Furthermore, if subjects are sensitive to statistically refined peer data, treatment participants who receive the high signal should choose to do more tasks than all baseline participants. If both hypotheses hold true, then treatment participants who receive the high signal should also choose to do more tasks than baseline participants who receive the high signal.

1.3 Results

Section 1.3.1 corroborates our theory's prediction that individuals will choose more costly actions to bolster self image in anticipation of social information. Section 1.3.2 supports the relevance of image signal extraction in our experiment.

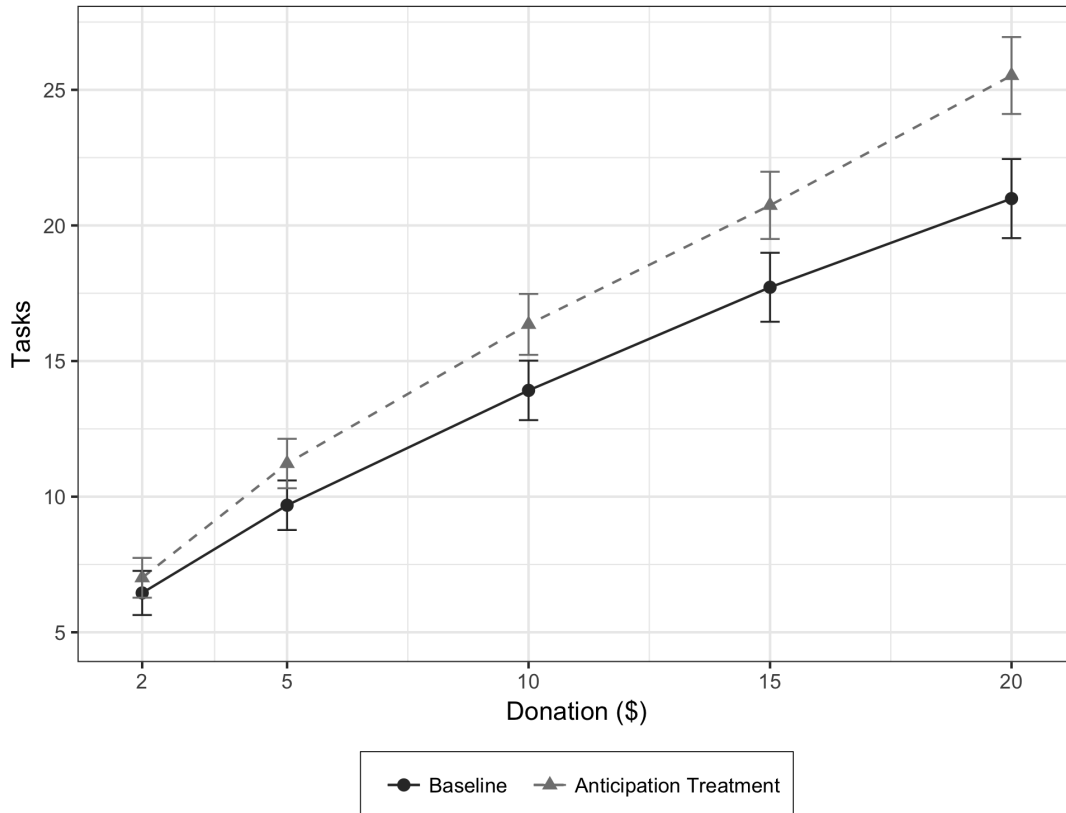
¹³The high signal – signal 1 – is the true signal.

¹⁴Subjects were not informed that they would be asked which signal they believed until after they made their decisions.

¹⁵Recall that all subjects were aware of both possible signals prior to drawing a signal.

1.3.1 Self-Image Results

Figure 1.5 plots the average task choice by donation amount of those in the baseline and anticipation treatments. Those in the anticipation treatment were, on average, willing to complete more captchas than those in the baseline. We also observe that task choice is increasing in the donation amount¹⁶ and that the treatment effect increases in the donation amount.



Notes: This figure plots the average number of tasks subjects were willing to complete for each possible donation. In concordance with self signaling, subjects choose to do more tasks if they anticipate learning how their choices compare to the choices of previous participants. This effect is increasing in the value of the gift to charity. Standard errors are shown in brackets.

Figure 1.5: Mean Task Choice by Donation

Table 1.1 presents these results via estimation of the following regressions:

$$y_{i,d} = \alpha_0 + \alpha_1 \text{Donation}_{i,d} + \alpha_2 \text{Treat}_i + \varepsilon_{i,d} \quad (1.1)$$

¹⁶94% of all participants displayed within-subject monotonicity in task choice across the donation amounts.

$$y_{i,d} = \beta_0 + \beta_1 \text{Donation}_{i,d} + \beta_2 \text{Treat}_i + \beta_3 \text{Treat}_i \times \text{Donation}_{i,d} + \varepsilon_{i,d} \quad (1.2)$$

where $y_{i,d}$ is the maximum number of tasks subject i was willing to complete for donation d . Regression (1.2) is more informative.¹⁷ From this regression we can see that for every additional \$5 donated to the ADRP, baseline participants were willing to do an additional four tasks while treatment participants were willing to do an additional five tasks. This result, significant at the 5% level, is useful for two principle reasons. First, strict monotonicity in task performance over donation size demonstrates the costliness of captcha transcription. Secondly, we find no evidence of any fixed treatment effect—the estimation of β_2 is indistinguishable from 0. Rather, the anticipatory image effect is entirely tied to the size of the gift given to the ADRP. This result is accommodated by our model. A nominal gift of \$2 gives little reason to hedge against bad news by making more costly choices. However, subjects choose to do 20% more tasks when the gift is \$15 or \$20. In other words, anticipatory image concerns only exist when sufficient charitable stakes are attached to an otherwise vacuous task.

1.3.2 Signal Extraction Results

In accordance with our theory, we find that participants’ choices are highly sensitive to the receipt of a statistically refined signal and unresponsive to a fully garbled signal. Recall that our theory requires that baseline choices should be independent of the obtained signal. Furthermore, treatment participants who receive the high signal should also choose to do more tasks than baseline participants who receive either signal. We focus on the high signal in treatment because it is the true signal and, therefore, the signal that we are statistically powered to evaluate. Figure

¹⁷Appendix Table A.1 confirms that the linearly interacted regression is the correct specification. The corresponding regression for Table A.1 is $y_{i,d} = \delta_0 + \sum_{d \in \{2,5,10,15,20\}} \delta_{1,d} 1_{i,donation=d} + \sum_{d \in \{2,5,10,15,20\}} \delta_{2,d} \text{Treat}_i \times 1_{i,donation=d} + \varepsilon_{i,d}$. From this regression we can not only see that task choices and the treatment effect are increasing in the donation, but also that a linear approximation of these effects fits well. For example, regression (1.2) predicts that the total average treatment effect for donations of \$15 and \$20 will be 3 and 4 tasks while the indicator regression finds the effects to be 3 and 4.5 tasks.

Table 1.1: Effect of Treatment on Tasks Allocation

	DV: Task Choice	
	(1)	(2)
Donation	0.903*** (0.045)	0.802*** (0.064)
Treatment	2.417* (1.452)	0.296 (1.152)
Treat \times Donation		0.204** (0.089)
Constant	4.367*** (0.928)	5.413*** (0.817)
R-Squared	0.208	0.211
Subjects	219	219
Observations	1095	1095

Notes: This table quantifies the results shown in Figure 1.5. Robust standard errors, clustered at the subject level, are presented in parentheses. There are 108 participants in the treatment and 111 in the baseline.

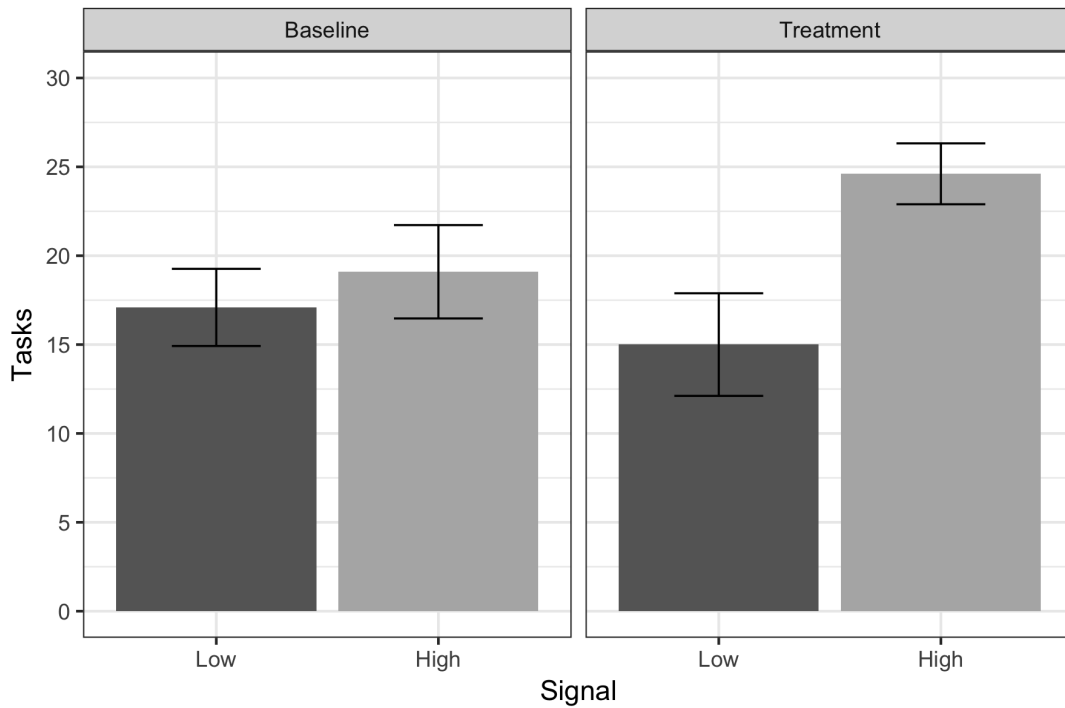
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

1.6 shows that baseline participant choices differ by a statistically insignificant two tasks between low and high signals. However, subjects who received the true signal with 15/16 probability chose, on average, 5.5 more tasks than those who received the high signal in baseline, and an average of 5.8 more tasks than all subjects in the baseline. These results are quantified in Table 1.2.

Additionally, we find that 53% of individuals believe signals received in baseline, while 87% of individuals believe signals received in treatment.¹⁸ Notably, individual beliefs are highly predictive of task choice. We find that those who believe the high signal is true are willing to do an average of 12 more tasks than those who believe the low signal. As appendix Table A.2 shows, this result holds true controlling for signal and treatment.

Lastly, we examine the distributional shift in task choice from baseline to treatment.

¹⁸This difference is significant at the 1% level. See appendix Table A.7. Beliefs by signal and treatment are shown in Figure 1.7.



Notes: This figure plots the average number of tasks subjects were willing to complete in exchange for a \$20 donation by treatment and signal. The high signal stated that more than 50% of previous participants were willing to complete at least 20 tasks while the low signal stated less than 25% were willing to complete 20 tasks. In concordance with image signal extraction, subjects chose to do significantly more tasks when they receive the high signal in treatment – where there is a 15/16 chance of receiving the true signal – than in the baseline – where there is an 8/16 chance of receiving the true signal. Brackets represent standard errors.

Figure 1.6: Average Tasks Choice by Signal, Treatment

The figures and tables we refer to in this analysis can be found in the appendix. Assuming monotonicity, Figure A.1 shows that the treatment effect is delivered by those who would have otherwise chosen to do fewer than 20 tasks in the baseline. This is reflected in Tables A.3 and A.4. Table A.3 demonstrates that 73% of treatment participants choose to do at least 20 tasks while only 52% of baseline participants choose to do 20 or more tasks. Further, quantile regression results in Tables A.4 and A.5 show that those below the median chose to do significantly more tasks in response to the treatment, while those above the median do not. Through the lens of our theory, these results, in conjunction with the results shown in Table A.2, suggest that those who choose to do fewer than 20 tasks in the baseline do so under the belief that most others made

the same decision. However, in the treatment, the refined signal corrects their beliefs and they choose to complete more captchas.¹⁹

Our theory relies on signal extraction as the mechanism underlying behavioral changes following the receipt of social information. Thus, establishing that signal extraction occurs in practice is vital to demonstrating the viability of our theory as predictive of behavior resulting from social interactions. Not only should the receipt of more reliable social information alter behavior, but it should alter individuals' underlying beliefs which induce this change in behavior. The results outlined above show that the receipt of a refined signal affects both beliefs and actions. Further, our results suggest that those individuals whose beliefs change are those that drive changes in average behavior across treatments.

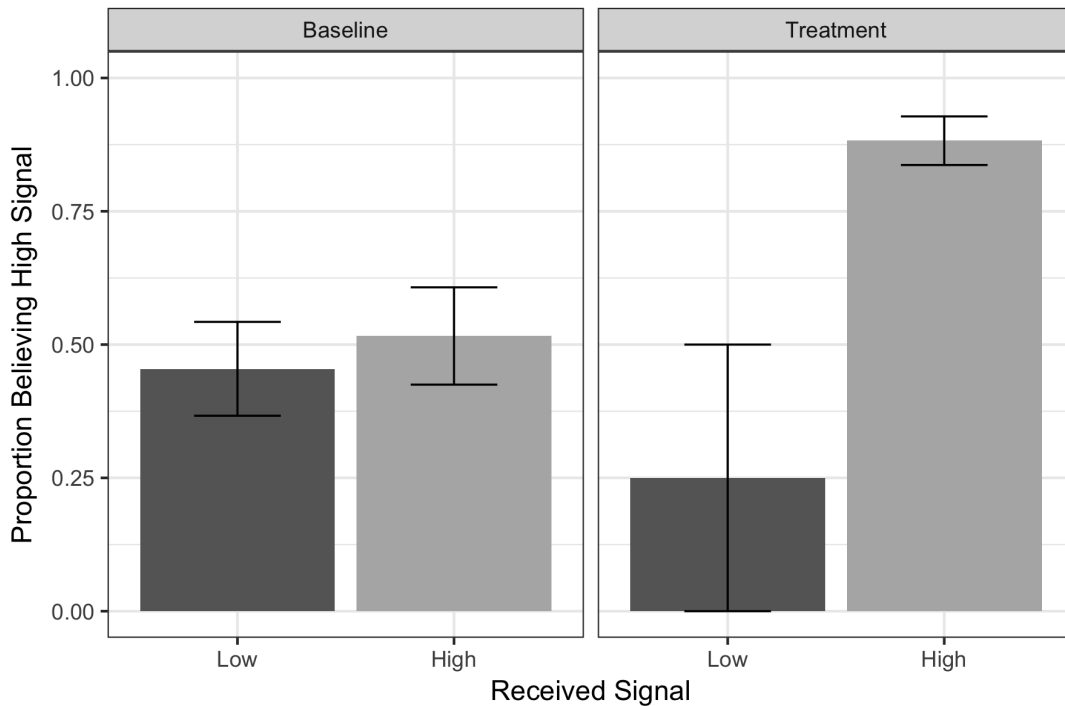
Table 1.2: Task Choice by Received Signal

	DV: Task Choice	
	(1) Both Signals	(2) High Signal
Treatment	5.847** (2.347)	5.511* (3.129)
Constant	18.062*** (1.686)	19.097*** (2.616)
R-Squared	0.050	0.040
Observations	119	82

Notes: Column 1 regresses task choice on treatment for all those in the experiment. Column 2 restricts the same regression to only those who received the high signal. Robust standard errors presented in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

¹⁹A plausible alternative story is that moral wiggle room/wishful thinking leads those who do fewer tasks to believe that everyone else also did few tasks. However, this story cannot explain why a more refined signal would correct deliberately chosen beliefs. Such a story would have to incorporate the dynamics of signal extraction wherein wishful thinking is easier to abide by when the received social information is less precise.



Notes: This figure plots the proportion of subjects who believed that more than 50% of previous participants were willing to complete at least 20 tasks by treatment and signal. In concordance with image signal extraction, participant beliefs are highly dependent on their signals in treatment, where there was a 15/16 chance of observing the true signal, and unresponsive in baseline where there was a 8/16 chance of observing the true signal. Brackets represent standard errors.

Figure 1.7: Distribution of Beliefs by Signal, Treatment

1.4 Conclusion

The economic importance of social interactions is well documented. However, there remain open questions at the core of understanding social pressure (Bursztyn and Jensen, 2017). This project complements existing theories of social conformity by theoretically and experimentally examining a justification for peer effects absent norms or stigmas. To that end, we demonstrate that if individuals rely on their own choices as well as social comparisons to form their self image, they will often mimic the behavior of their peers. With no norm to follow, (partial) social conformity arises if decision makers are more afraid of learning that they are of low social type than they are eager to prove that they are of high social type.

We are able to empirically distinguish our rationalization of peer effects from others through novel a prediction on the anticipation of social information. Our model predicts that individuals will make more costly and image-positive choices in anticipation of social information. We test this prediction in a series of lab experiments. In accordance with our theory, experiment participants choose costlier actions in anticipation of social information.

Lastly, a potentially relevant literature documents an important relationship between uncertainty and self-serving preferences (Dana, Weber and Kuang, 2007; Andreoni and Bernheim, 2009; Exley, 2016). Excuse-driven preferences may exacerbate the effect of perceived signal interference in image signaling. Therefore, exploring this relationship may provide better predictive power over the environmental conditions and policy interventions that will catalyze or mitigate social conformity. Lastly, our theory can be naturally extended to yield predictions on when individuals will seek or avoid information. Such predictions could potentially be used to understand when social information is deemed harmful, beneficial, or irrelevant to one's own image. Such a research agenda has potentially broad implications for work-place, school, and social practices and may provide greater predictive power on how to shape peer influence across diverse social contexts.

Acknowledgements

Chapter 1, in full, is currently being prepared for submission for publication of the material. Martinez, Seung-Keun; Bigenho, Jason. "Social Comparisons in Peer Effects." The dissertation author was one of two primary investigators and authors of this material.

Chapter 2

Procrastination in the Field: Evidence from Tax-Filing

2.1 Introduction

Present-biased preferences (Laibson, 1997; O'Donoghue and Rabin, 1999, 2001) are critical for understanding deviations from the neo-classical benchmark of exponentially discounted utility (Samuelson, 1937; Koopmans, 1960). Prominent anomalies such as self-control problems, the demand for commitment devices, and procrastination in task performance revolve around the tension between long-term plans and short-term temptations inherent to these models.

Identifying present-biased preferences from field data faces a natural challenge. Though the forces of short-term temptations may be observed in behavior, researchers will rarely have access to data on long-term plans. It is potentially for this reason that the body of evidence in support of present-biased preferences comes largely from laboratory study.¹ When field data are used, evidence of present bias is generally calibrational, based on behavior alone; suggesting

¹Frederick, Loewenstein and O'Donoghue (2002) provide a detailed review of the experimental literature. The overwhelming majority of such studies employ choices over time-dated monetary payments. There exist important critiques of this literature related to the fungibility of money. Only a few studies employ measures of consumption to investigate present bias. See Sprenger (2015) for recent discussion of this literature.

either that implausibly high levels of exponential discounting would be required to rationalize an observed data set (see, e.g., Fang and Silverman, 2009; Shapiro, 2005); or that exponential discounting provides substantially worse fit than a present-biased alternative (see, e.g., Laibson et al., 2017). One notable exception to this tradition is DellaVigna and Malmendier (2006), who use the distinction between reported plans and observed gym attendance to make inference on present bias.²

This paper explores present-biased preferences in field data focusing on the often-discussed problem of procrastination in tax-filing (Slemrod et al., 1997; O’Donoghue and Rabin, 1999).³ Linking techniques from structural estimation of dynamic discrete choice (Hotz and Miller, 1993; Arcidiciano and Ellickson, 2011) to empirical strategies from public finance (Chetty et al., 2011), we propose a ‘missing mass’ method identifying deviations from exponential discounting. The counterfactual plan of action under exponential discounting is constructed from the dynamic discrete choice model and contrasted with true behavior, delivering the critical comparison between plan and behavior. Present-biased procrastination predicts that true filing close to the tax deadline will exceed counterfactual exponential filing.

Differentiating procrastination from optimal delay is notoriously difficult in settings like tax filing. First, if costs of filing are stochastic, one should expect to see increased filing close to the deadline as the option value of future filing diminishes (Slemrod et al., 1997).⁴ Hence, increased filing close to the deadline is not sufficient to identify procrastination, a point which

²It should be noted, however, that reported plans in DellaVigna and Malmendier (2006) come from a different sample of individuals than most of their observed behavior. Data on reported plans suffer from a general problem of incentive compatibility: the researcher cannot know if a self-report is what the agent truly believes, and if she attempts to incentivize this report, then the agent may appear to follow through on their plans only to collect the incentive. Perhaps for this reason a strand of literature has developed sidestepping the necessity of having both plan and behavior by providing smoking-gun evidence of sophisticated present bias in the form of commitment demand (see, e.g., Ashraf, Karlan and Yin, 2006; Ariely and Wertenbroch, 2002; Bisin and Hyndman, 2014; Kaur, Kremer and Mullainathan, 2010; Gine, Karlan and Zinman, 2010; Mahajan and Tarozzi, 2011). Laibson (2015) provides a recent discussion on the calibrational plausibility of commitment demand in the presence of uncertainty and commitment costs, indicating that commitment may be the exception rather than the rule in many settings.

³Slemrod et al. (1997) coined the phrase ‘April 15th Syndrome’. O’Donoghue and Rabin (1999) use tax filing as their leading example of procrastination predicted by dynamic inconsistency.

⁴Slemrod et al. (1997) explicitly notes the potential importance of stochastic costs in rationalizing the observed distribution of filing behavior including both heterogeneity and late filing. Though the authors term late filing behavior ‘procrastination’ they note explicitly that their rationalization is dynamically consistent.

calls into question a number of studies that infer present bias from delay alone (e.g., Brown and Previtro, 2014; Brown, Farrell and Weisbenner, 2016; Frakes and Wasserman, 2016). This project overcomes this issue by explicitly recognizing stochastic costs in the construction of both our exponential benchmark and our present-biased alternative. Second, behavioral models of present bias make use of additional parameters beyond the exponential discounting formulation. Hence, improved in-sample fit for the behavioral model should largely be expected, precluding strong inference on this basis. This project overcomes this issue by linking our estimates with both responses to changing filing incentives and experimental measures of time preferences to provide out-of-sample tests for our estimated models. Third, in the specific context of tax-filing, Internal Revenue Service data generally only provides the date which the tax return is processed and not the date of filing (Slemrod et al., 1997; Benzarti, 2015), generating a non-standard measurement error problem for the timing of behavior.⁵ This project overcomes this issue by using precise data on tax return initiation.

In a sample of 22,526 low-income tax filers in the City of Boston from 2005 to 2008, we identify a substantial missing mass in filing behavior relative to an exponential benchmark. Despite a high degree of estimated impatience under exponential discounting, we document a wide deviation between actual and predicted filing probabilities as the end of the tax season approaches. These deviations deliver a missing mass of around 80% additional tax filers relative to the exponential benchmark in the last seven days prior to the deadline.

We interpret our missing mass as evidence of present-biased procrastination in filing and

⁵Slemrod et al. (1997) use the 1998 Internal Revenue Service Individual Master File of 95,000 tax returns for the 1988 tax year appended with the date assigned by the IRS Service Center upon receipt. Reference is subsequently made to the date of processing throughout the text and more than 25% of returns occur in the second half of April (April 17th was the deadline in 1989). Note is made of the potential for IRS delays in assigning dates for returns received during ‘the last-minute surge of filings in April’ (p.698). Slemrod et al. (1997) make use of the 1979-1988 Statistics of Income Panel to conduct longitudinal analysis. The process date is not available in this panel so they use the week at which the return was posted to the IRS Individual Master File and note that the median posting is beyond the tax filing deadline, but that substantial correlation exists between the posting weeks and the process dates. Benzarti (2015) also makes use of the Statistics of Income Panel posting dates and relates them to itemized deductions. As alluded to by Slemrod et al. (1997), a non-standard measurement error problem may be generated if one wishes to use IRS process or posting dates as a proxy for filing times. The correlation between filing dates and process dates is likely influenced by the number of filers. Hence, the concordance between the true measure and the proxy changes through time and is likely worst close to the deadline.

bolster this interpretation with three pieces of evidence. First, the data can be rationalized with a relatively small degree of present bias close to recent empirical estimates. This gives calibrational support to our interpretation. Second, within our sample period lies the 2008 Economic Stimulus Act, which generated plausibly exogenous variation in filing benefits for 2008 Stimulus Payment recipients.⁶ If individuals were as impatient as our exponential estimates imply, they should exhibit no response to these additional filing benefits.⁷ In contrast, difference-in-difference estimates suggest stimulus recipients file around 2 days earlier under the stimulus, a sensitivity that is well predicted out-of-sample by our estimated degree of present bias. Third, we have access to a sub-sample of 1114 individuals who completed incentivized time preference experiments in 2007 and 2008. This sample allows us to link experimental measures of present bias to deviations from exponential discounting in tax filing behavior. The gap between exponentially-predicted and actual filing behavior correlates strongly with our experimental measures. Present-biased subjects file disproportionately later than exponential prediction relative to other experimental subjects, and this difference is most pronounced towards the end of the tax filing season.

The methods implemented and the results obtained in this paper contribute to several strands of literature in behavioral economics and the broader field.

First, our methods identify behavioral anomalies in field behavior by relying on structural estimation of a specific, neoclassical model to construct the counterfactual benchmark. A growing body of empirical projects identify behavioral forces using measures of missing mass relative to a traditional atheoretic benchmark such as a smooth or unchanging distribution of behavior (see, e.g., Rees-Jones, 2013; Benzarti, 2015; Allen et al., Forthcoming). Use of such minimally parametric counterfactuals is ideal for examining the adherence of behavior to a class of similar smooth theories or where the benchmark theory requires only such minimal restrictions.

⁶It should be noted that the timing of the 2008 Stimulus Payments did not depend on the timing of tax-filing. However, in the advertisement of the program, the IRS clearly conveyed linkages between tax filing timing and Stimulus Payment receipt. See section 2.2 for discussion and details.

⁷The 2008 Economic Stimulus Act generated additional payments on the order of \$300- \$600 for many families. The IRS conveyed that these payments would be received several weeks after the filers normal refund. With our estimated extreme impatience such additional later benefits are discounted almost to zero. See section 4.3 for further detail.

In many cases, however, researchers may be interested in using missing mass methods to reject a single model of behavior rather than a class thereof, or generating a test of a more nuanced model prediction. Using structural techniques to guide counterfactual construction can help to generate tight tests of underlying models in such cases. The specific dynamic discrete choice techniques implemented for our problem are readily portable to other intertemporal problems where researchers may be interested in potential deviations from exponential discounting.⁸

Second, and related to the point above, once a single counterfactual model of behavior is rejected, many candidate theories may arise to ‘rationalize’ the missing mass. As carefully demonstrated by Einav, Finkelstein and Schrimpf (2016) for the case of bunching estimators, different candidate theories could make dramatically different predictions for responsiveness to key policy variables. Assessing the predictive validity of our favored candidate theory by examining responsiveness to the exogenous changes in filing incentives induced by the 2008 Stimulus Payments is a key contribution of this paper. Such out-of-sample steps are particularly important to take if candidate theories are behavioral, as appeal to additional free parameters in behavioral models will generally deliver greater in-sample fit. We are aware of no such out-of-sample tests in the prior literature for present-biased preferences. DellaVigna et al. (Forthcoming) provides one recent demonstration of the value of such exercises for behavioral models of reference dependence in the context of job search.⁹

⁸One recent example of such a setting is Heffetz, O’Donoghue and Schneider (2016), who investigate delay in payment of parking tickets in New York. The authors demonstrate a sharp increase in the hazard rate of payment around a first penalty deadline. They also show a sensitivity of behavior to changed notification policy. The authors interpret the data as evidence of forgetting, and note the plausibility of present bias but the challenge of identifying such time preferences in their setting. Our method for estimating present bias outlined in section 4.2 (with some adjustments for multiple deadlines and assumptions about the role of notifications) may be helpfully applied. Additionally, projects such as Brown and Previtro (2014); Brown, Farrell and Weisbenner (2016) and Frakes and Wasserman (2016), which infer present bias from delay alone, could potentially yield sharper conclusions with the implementation of such methods.

⁹DellaVigna et al. (Forthcoming) examine exit from unemployment under reference-dependent preferences and show their preferred model not only rationalizes job finding hazard rates better than a standard model but also provides improved prediction for responsiveness to changes in the structure of unemployment insurance benefits. DellaVigna et al. (Forthcoming) also include a second behavioral parameter beyond reference dependence in the form of present bias. Their objective in doing so is to consider a more broad model accounting both for patterns of asset accumulation and job search. Present bias in their setting acts in a very similar way to extreme impatience, and hence, serves primarily a calibrational role. However, the outlined estimation strategy is of general value. Indeed, their assumption full naivete for this exercise inspired our own. See section 4.2 for detail.

Third, we use experimental measures of time preferences to validate our interpretation of present bias in tax filing. Importantly, our experimental measures come from choices over time-dated monetary payments. A growing discussion in the behavioral literature has questioned the use of such measures given the fungibility of money (Cubitt and Read, 2007; Chabris, Laibson and Schuldt, 2008; Andreoni and Sprenger, 2012; Augenblick, Niederle and Sprenger, 2015; Carvalho, Meier and Wang, 20016; Dean and Sautmann, 2016). Given the low income of our sample and the plausibility of liquidity constraints, in our setting there may be reason to believe experimental responses are driven by true preferences rather than arbitrage.¹⁰ Our findings demonstrate such plausible informativeness given that our measures relate to apparent procrastination in filing behavior. This also complements the recent contributions of Mahajan and Tarozzi (2011), who show that experimentally elicited preference measures can be productively incorporated into structural approaches in this domain.¹¹

Fourth, our paper delivers potentially policy relevant measures of present bias for the population in question. With credible estimates of such preferences, we can identify not only those individuals who are likely to have procrastinated in their tax filing, but also conduct welfare analysis for a number of potential policy interventions. Among these analyses is an evaluation of the recently enacted Protecting Americans from Tax Hikes (PATH) Act of 2015, which embargoed the *entire* refund of Earned Income Tax Credit recipients until February 15th for the 2017 tax filing season. We evaluate such measures under both exponential discounting and present-biased preferences, providing an input to evaluating elements of tax policy that have,

¹⁰Our data give us some opportunity to examine the marginal propensity to consume (MPC) out of expected income. In survey response, we find an average MPC of around 0.75. This relatively high MPC echoes findings such as Souleles (1999), who documents a quarterly MPC out of refunds of around 0.65 on average, though largely spent on durables. Our subjects overwhelmingly report intending to spend on bill and debt payment as well as non-durables, with only fifteen percent of respondents report intentions of spending in such categories as paying for a home, car, home-improvements or school. See Table 1 for detail.

¹¹Mahajan and Tarozzi (2011) incorporate experimental measures for time inconsistency, beliefs about disease infection, and purchase and treatment decisions for insecticide treated bednets to estimate the extent of present bias and ‘sophistication’ thereof. In their setting the experimental measures are used directly in estimation, while in ours they are used for validation ex-post. One point noted by Mahajan and Tarozzi (2011) is that in their case the experimental measures themselves wind up having limited predictive power for estimates of present bias that result from their structural exercise.

to our knowledge, not yet received attention.

The paper proceeds as follows: section 2 describes the data and the experimental procedures. Section 3 then presents our empirical design and the construction of our missing mass measures. Section 4 presents results, and section 5 concludes.

2.2 Data

Our exercise makes use of three key data sources: 1) tax filing data from the City of Boston, Massachusetts; 2) variation in tax refunds due to the 2008 Stimulus Act; and 3) experimental measures of time preferences from tax filers in 2007 and 2008.

2.2.1 Tax Filing Data

The data in this paper comes from 22 Volunteer Income Tax Assistance (VITA) sites in Boston, Massachusetts from the years 2005 to 2008. VITA sites are organized by the City of Boston and the Boston Tax Help Coalition. VITA sites provide free tax preparation assistance to low-to-moderate income households in specific neighborhoods in order to help them claim valuable tax credits such as the Earned Income Tax Credit (EITC). Boston's VITA sites began in 2001 and continue to present. As of the 2015 tax filing season there were 27 VITA sites in operation around the city, processing a total of 12,940 returns and securing around \$22.8 million in refunds of which \$8.6 million were EITC payments (see bostontaxhelp.org for current information and details).

VITA sites generally open in mid-January and close at the tax filing deadline around April 15. Most sites have specific days and hours of operation, though some are open by appointment only. Potential filers are encouraged to bring all required documentation (photo ID, W2 forms, 1099 forms, etc.) to the VITA site. Sites have an in-take coordinator who provides filers with a check-list of required documents. Filers are usually processed on a first-come, first-served

basis and waiting times at popular sites can be substantial during busy periods. Upon reaching the front of the queue, the tax-filer meets with a volunteer preparer, the return is entered, and subsequently filed with the Internal Revenue Service electronically.

In 2005 and 2006 we have access to the date the return was electronically filed, while in 2007 and 2008 we have both the date the return was initiated and the date the return was filed. Around 80% of returns in our final sample are filed within two days of initiation in 2007 and 2008, delivering a close correspondence between when returns are initiated and when they are electronically sent to the IRS. As the deadline approaches, the correspondence grows, with around 90% of returns filed within two days of initiation during the last week of the tax season.¹² Critical for the present study, we are able to observe the full return information including the date each return was initiated and/or electronically filed, whether any refund would be received via direct deposit or paper check, and the size of federal refunds.

From 2005 to 2008 a total of 32,641 tax returns were initiated at VITA sites. Of these, 26,040 (87.6%) were filed electronically with documented acceptance by the IRS in our data.¹³ To have the most precise measure possible for when individuals decide to file, we use the electronic filing date from 2005-2006 and the initiation date from 2007-2008 as our measured filing date. We recognize that the 2005-2006 filing times may slightly overstate the timing of tax filing relative to 2007-2008, and so also provide all estimates using only the latter years for which precisely measured initiation data are available¹⁴.

We restrict the sample along several other dimensions for our study of procrastination. First, we remove 212 (0.8%) individuals who have filing dates after the filing deadline. Second, we focus on only the 11 weeks prior to the deadline such that the majority of subjects can be

¹²The mean (median) filing lag for 2007 and 2008 is 2.3 (0) days, and in the last week of the tax season the mean (median) filing lag is 0.8 (0) days.

¹³28,606 returns were ever sent to the IRS. Of these 2215 (7.7%) returns have a documented rejection. Though many of these returns were subsequently filed and accepted, our data have the initial electronic filing date overwritten by the subsequent filing. For a further 352 (1.2%) returns, we have no documented acceptance. We are hesitant to use these data as we are unsure either of the first filing date or of when, if ever, a refund was received.

¹⁴To this end, we reproduce our primary estimation results, shown in table 2.2, with only 2007 data. The results can be found in appendix table B.2

expected to have received their primary tax documents such as W2s.¹⁵ This eliminates 621 (2.4%) observations. Third, we focus only on subjects with weakly positive refunds, eliminating a further 1457 (5.6%) filers.¹⁶ Fourth, we eliminate 1174 (4.5%) individuals with zero dollars of taxable income and zero dollars of refund. Such individuals would not generally need to file taxes, but do so likely because of the 2008 Stimulus Payments which provided rebates to such filers.¹⁷ Fifth, a small number of subjects, 65 in total, appear to file on Sundays when VITA sites are generally closed and no electronic filing should be possible. We believe these special cases correspond to appointments or VITA site workers filing on their own behalf and, hence, drop these observations as well. In total, these restrictions eliminate 13.5% of observations leaving a usable sample of 22,526 individual tax filings.

Table 2.1, Panel A presents summary statistics for our sample across the four years of our study. Tax filers are around 38 years old, earn around \$17,000 in adjusted gross income, and receive sizable federal refunds of around \$1,400. Tax filers have slightly more than half a dependent, and around 10% of subjects receive unemployment benefits in any given year. In addition to the above measured socio-demographics which are captured directly from tax returns, VITA sites also ask tax-filers to complete a socio-demographic survey to identify gender, race, and education levels. Response rates for these questions vary from 77% to 78%. Panel B demonstrates that conditional on responding, the majority of tax filers report they are female, African-American, and without college experience.

Table 2.1, Panel C presents two important time related variables: the number of days until the filing deadline and whether or not a tax-filer opts to receive their refund by direct deposit. In order to identify the number of days until the tax filing deadline we subtract the deadline

¹⁵Our empirical exercise will require individuals to project whether individuals with similar characteristics will file in the next period. Including the relatively sparse data outside of 11 weeks, generates some missing or extreme projections.

¹⁶Because tax filers have relatively low incomes, they generally receive substantial proportional refunds associated with the EITC and other tax credits. Our empirical exercise estimates an optimal stopping problem with costs of filing and refund benefits. We do not explicitly model the kink in incentives associated with filing beyond the deadline and incurring a tax penalty if one has a negative refund. With negative refunds, individuals will have incentives primarily to file close to the deadline.

¹⁷Indeed, 906 of 1174 (77.2%) individuals who fall into this category are observed in 2008.

date from the filing date. On average filing occurs 44 days before the tax filing deadline with substantial heterogeneity. Individuals who receive their refund by direct deposit can expect to receive their refund substantially earlier than those who do not. Given that only 40% of our sample opts for direct deposit, this presents potentially important cross-sectional variation in the timing of refund receipts.¹⁸

Our estimation strategy attempts to estimate preferences from the timing of tax filing and refund receipt. A critical question is what tax refunds are used for. Prior estimates of marginal propensity to consume (MPC) from tax refunds show that as much as 65% of each dollar of refund is consumed within a quarter, and, for liquidity constrained individuals, refund receipt leads to substantial increases in non-durable consumption (Souleles, 1999). Recognition of such high MPCs and the plausibility of liquidity constraints is important in our sample as our baseline estimation strategy implicitly assumes that the timing of refund receipts perfectly correlates with the timing of consumption.¹⁹ In Table 2.1, Panel D we examine the self-reported intentions of refund use for responding individuals. First, individuals were asked which of 16 non-exclusive categories their refund would be used for.²⁰ Eighty percent of respondents who reported at least one category reported intending to use the refund to pay bills or debts²¹, and twenty-eight percent of respondents reported a non-durable category such as buying groceries or going shopping²². Fifteen percent of subjects reported a durable category such as purchasing a car or home²³. Twenty-four percent of subjects reported a savings category such as saving for a home or a rainy

¹⁸Each year (through 2012) the IRS provided tax-filers with a refund cycle linking electronic filing and acceptance dates to dates when direct deposits would be sent and paper checks would be mailed. In general, accepted returns are batched by week and paper checks are mailed one week after direct deposits are sent. For our baseline estimate we ignore the discontinuities in refund receipt induced by this batching protocol as it is unlikely that tax-filers in our sample would have access to such information. Additionally, the data do not appear to reflect the batch discontinuities with individuals bunching close to batch endpoints.

¹⁹In section 4.1.1 we relax this assumption, assuming the refund is consumed over a one-month horizon.

²⁰The categories were: 'Buy Groceries', 'Pay Bills', 'Pay Back Debts', 'Pay Old Taxes', 'Go on Vacation', 'Go Shopping', 'Buy a Home', 'Save for a Home', 'Pay Medical Bills', 'Buy a Car', 'Save for a Car', 'Pay Child Expenses', 'Pay for School', 'Save for School', 'Save for a Rainy Day', and 'Home Improvement'.

²¹The relevant choices in this category are 'Pay Bills', 'Pay Back Debts', 'Pay Old Taxes', 'Pay Medical Bills', and 'Pay Child Expenses'.

²²The relevant choices in this category are 'Buy Groceries', 'Go on Vacation', and 'Go Shopping'.

²³The relevant choices in this category are 'Buy a Home', 'Buy a Car', 'Pay for School', and 'Home Improvement'.

day²⁴. That the majority of refunds were intended for consumption is echoed in respondents self-reported savings intentions. Individuals were asked what percentage of their refund they intended to save in one of five categories.²⁵ The mean self-reported savings percentage is around 24%, indicating an intended MPC out of refund of around 0.75. Further, that 80% of respondents report intending to use refunds for payment of bills or debts indicates limited access to liquidity.

2.2.2 Economic Stimulus Act 2008

A key component of our exercise attempts to predict sensitivity of filing behavior to the exogenous change in filing incentives provided by the Economic Stimulus Act of 2008. Under the Economic Stimulus Act of 2008 (H.R. 5140) passed on February 7, 2008, tax filers earning less than \$75,000 (\$150,000 for joint filers) received ‘Recovery Rebates’ between \$300 and \$1200, depending on filing status and income levels. The Stimulus Payments were announced in February 2008. In practice, these payments were generally disbursed between late April and July of 2008 depending on the social security number of the tax filer. In 2008, 90% of the filers we observe qualified to receive Stimulus Payments. Appendix Figure B.1 presents the histogram of Stimulus Payments calculated from individual tax return data.²⁶

Prima-facie, the 2008 Stimulus Payments, whose values were based on predetermined income and demographics, could provide for exogenous variation in refund sizes and give potential for difference-in-difference investigation across recipients and non-recipients. However, the timing of tax-filing had no true impact on the receipt of the Stimulus Payments and hence did not truly influence the intertemporal tradeoffs. Nonetheless, it is not clear that tax filers

²⁴The relevant choices in this category are ‘Save for a Home’, ‘Save for a Car’, ‘Save for School’, and ‘Save for a Rainy Day’.

²⁵The categories were: ‘0%’, ‘<10%’, ‘10-25%’, ‘25-50%’, and ‘>50%’. We take the midpoint of the interval implied by response (bottom coded for ‘0%’ responses), as the self-reported savings intention.

²⁶The 2008 stimulus rebate took two forms. The first was allocated according to filing status, tax liability, adjusted gross income, and number of dependents. The second was allocated according to filing status, number of dependents, adjusted gross income, social security, and other qualifying income. The first type was phased in and out according to AGI. Each individual received the larger of the two rebates. The exact formula is detailed in the ‘‘Technical Explanation of the Revenue Provisions of H.R. 5140.’’ Using each individual’s 1040A data we calculated their rebates with python script.

Table 2.1: Summary Statistics

<i>Panel A: Tax Return Information</i>			<i>Panel B: Demographics:</i>			<i>Panel C: Filing and Direct Deposit</i>			<i>Panel D: Intended Refund Uses</i>		
Variable	Obs	Mean (s.d.)	Variable	Obs	Mean (s.d.)	Variable	Obs	Mean (s.d.)	Variable	Obs	Mean (s.d.)
Age	22524	37.87 (15.54)	Female (=1)	17541	0.648 (0.48)	Days until Deadline	22526	44.06 (21.68)	Intended Savings (%)	9303	23.96 (26.85)
Adjusted Gross Income	22526	16924.8 (13367.2)	Black (=1)	17268	0.560 (0.50)	Direct Deposit (=1)	22526	0.397 (0.49)	Durable	11107	0.15 (0.36)
Federal Refund	22526	1419.82 (1618.2)	College Experience (=1)	16757	0.155 (0.36)				Non-Durable	11107	0.28 (0.45)
# Dependents	22526	0.535 (0.87)							Bills and Debt	11107	0.79 (0.41)
Unemployed (=1)	22526	0.095 (0.29)							Saving	11107	0.24 (0.43)

at VITA sites, or anyone else for that matter, were aware of this point. The initial February 2008 announcement did not clarify that the timing of Stimulus Payments was decoupled from dates of tax-filing, but noted only that the payments would begin being made in May of 2008.²⁷ Furthermore, the IRS' documentation of the Stimulus Payments may have created the impression that Stimulus Payment timing was linked to filing dates. Filers examining the Frequently Asked Questions website related to the Stimulus Payment asking 'When will I receive my Stimulus Payment?' were told²⁸:

Processing times for tax returns and Stimulus Payments vary. If you are getting a regular tax refund, the IRS will send you that refund first. Normally, your Stimulus Payment will follow one to two weeks later.

Such information likely gave filers the impression of a tight link between filing times and Stimulus Payment receipt. Hence, the 2008 Stimulus Payments generated plausibly exogenous variation in the benefits of filing. In Section 2.4.2 we use the 2008 Stimulus Payment data and attempt to predict responsiveness to these changing filing incentives under our estimated preferences out-of-sample.

2.2.3 Experimentally Elicited Time Preferences

For a subsample of tax filers, we have independent measures of time preferences elicited using experimental methods. In 2007 and 2008 at one VITA site, in Roxbury, MA, we conducted incentivized intertemporal choice experiments throughout the tax filing season. These data are discussed in detail in Meier and Sprenger (2015), which analyzes stability of elicited preferences at the aggregate and individual level.²⁹ A total of 1902 individuals in our sample received

²⁷Appendix B.5 reproduces the IRS announcement. Additionally, the actual process of payment was not described in the technical description of the Stimulus Payments. Appendix B.6 reproduces the relevant portions of the technical explanation of the revenue provisions of the Economic Stimulus Act of 2008.

²⁸See <https://www.irs.gov/uac/Economic-Stimulus-Payment-Q&As:-When-Will-I-Get-the-Payment%3F> for details. This website was updated in July 2008 likely to reflect the volume of payment to date.

²⁹Meier and Sprenger (2015) demonstrate stable choice profiles and corresponding parameter estimates at both levels and a one-year correlation in behavior of around 0.5. Instability in experimental choice is largely orthogonal from demographics or changes in financial situation.

tax-filing assistance in 2007 and 2008 at the Roxbury VITA site. Of these, 1,794 filed their taxes on one of the days the experiment was conducted and were eligible to participate. In both years VITA site intake material included identical, incentive-compatible choice experiments to elicit time preferences. The choice experiments were presented on a single colored sheet of paper and were turned in at the end of tax-filing for potential payments (see below). The experimental paradigm is presented as Appendix B.7. 1296 individuals, (72.5%) elected to participate. Appendix Table B.1 presents observable characteristics of our experimental subjects and compares them to the observables of non-participating subjects at the Roxbury VITA site. Experimental subjects appear similar on observables to non-participating subjects.

Individual time preferences are elicited using identical incentivized multiple price lists in both 2007 and 2008 (for similar approaches to elicit time preferences, see Coller and Williams, 1999; Harrison, Lau and Williams, 2002; McClure et al., 2004; Dohmen et al., 2006; Tanaka, Camerer and Nguyen, 2010; Burks et al., 2009; Benjamin, Choi and Strickland, 2010; Ifcher and Zarghamee, 2011). Individuals were given three multiple price lists and asked to make a total of 22 choices between a smaller reward, X , in period t and a larger reward, $Y > X$, in period $t + \tau > t$. We keep Y constant at \$50 and vary X from \$49 to \$14 in three time frames. In Time Frame 1, t is the present, $t = 0$, and τ is one month. In Time Frame 2, t is the present, and τ is six months. In Time Frame 3, t is six months from the study date, and τ is again one month. The order of the three time frames was randomized. Appendix B.7 provides the full set of choices.

In order to provide an incentive for truthful choice, 10 percent of individuals were randomly paid one of their 22 choices (for comparable methodologies and discussions, see, e.g., Harrison, Lau and Williams, 2002). This was done with a raffle ticket, which subjects took at the end of their tax filing and which indicated which choice, if any, would result in payment. To ensure credibility of the payments, we filled out money orders for the winning amounts on the spot in the presence of the participants, put them in labeled, pre-stamped envelopes and sealed the envelopes. The payment was guaranteed by the Federal Reserve Bank of Boston

and individuals were informed that they could always return to the head of the VITA site (the community center director) where the experiment was run to report any problems receiving the payments. Money orders were sent by mail to the winner's home address on the same day as the experiment if $t = 0$, or in one, six, or seven months, depending on the winner's choice. All payments were sent by mail to equate the transaction costs of sooner and later payments. The details of the payment procedure of the choice experiments were kept the same in the two years and participants were fully informed about the method of payment.

The multiple price list design yields 22 individual-level decisions between smaller, sooner payments X and larger, later payments Y . We term the series of decisions between X and Y a *choice profile*. We make one restriction on admissible choice profiles: that the choices satisfy monotonicity within a price list.³⁰ Roughly 86% of our sample, or 1114 individuals satisfy this restriction

In order to identify present bias from the observed choice profiles, we examine choices made in Time Frame 1 ($t = 0, \tau = 1$), and Time Frame 3 ($t = 6, \tau = 1$). Let X_1^* be the smallest value of X for which an individual chooses X over Y in Time Frame 1, and let X_3^* be the smallest value of X for which an individual chooses X over Y in Time Frame 3. An individual is coded as *Present-Biased* if $X_1^* < X_3^*$, having expressed more patience over six vs. seven months than over today vs. one month. Similar measures for identifying time preferences from experimental data have been employed by Coller and Williams (1999); Harrison et al. (2005); McClure et al. (2004); Dohmen et al. (2006); Tanaka, Camerer and Nguyen (2010); Burks et al. (2009); Benjamin, Choi and Strickland (2010); Ifcher and Zarghamee (2011); Meier and Sprenger (2010). Of our 1114 subjects, 360 (32%) are classified as *Present-Biased*.³¹

³⁰That is, individuals do not choose X over Y and Y over X' if $X' < X$. This restriction is equivalent to focusing on individuals with unique monotonic switch points and individuals without any switch points in each price list. The level of non-monotonicity obtained in our data compares favorably to the level obtained in other multiple price list experiments with college students, where around 10% of individuals have non-unique switch points (Holt and Laury, 2002) and is substantially below some field observations where as many as 50% of individuals exhibit non-unique switch points (Jacobson and Petrie, 2009). For non-monotonic subjects we are unable to have a complete record of their choices as we measure only their first switch point and whether they switched more than once. Price list analysis often either enforces a single switch point (Harrison et al., 2005) or eliminates such observations.

³¹111 (10%) of subjects are classified as *Future-Biased* with $X_1^* > X_3^*$. The remaining 643 subjects (58%) exhibit

In Section 2.4.3, we link this experimental measure of present bias to deviations from exponential discounting in tax filing behavior.

2.3 Empirical Strategy

We construct and estimate optimal stopping models for the timing of tax filing based on the techniques from dynamic discrete choice. Estimates are generated for both exponential discounting and quasi-hyperbolic discounting and counterfactual distributions for predicted filing patterns can be contrasted with actual filing behavior. Our methodology and notation borrows heavily from the formulations of Hotz and Miller (1993); Arcidiciano and Ellickson (2011). For space considerations, Appendix B.2 provides a more detailed presentation than that provided here.

2.3.1 Tax Filing as Exponential Dynamic Discrete Choice

An individual's decision to file taxes can be viewed as an optimal stopping problem. In each period before the filing deadline, the individual decides whether to incur a realized cost and receive the benefit of sooner receipt of their refund, or to wait to file on a future date. Though all individuals in our sample receive positive refunds, and hence face no penalty for late filing, we assume the costs of filing become sufficiently high once the VITA sites close such that no individual ever desires or forecasts filing after the deadline.³²

There are N individual tax filers, indexed by i . Time is discrete, indexed by t , with T denoting the period of the tax deadline. In each period, tax filers take actions a_{it} . They either decide to postpone filing ($a_{it} = 0$) or to file ($a_{it} = 1$). Let f_{it} denote the individual's filing status in period t such that $f_{it} = 1$ if the individual has not yet filed by period $t - 1$ and $f_{it} = 0$ if the individual has filed by period $t - 1$. We assume that each individual will receive a positive refund,

$X_1^* = X_3^*$ consistent with exponential discounting.

³²In principle, individuals with positive refunds have three years to file and claim their refunds from the IRS. After three years, the funds become the property of the United States Treasury.

b_i , constant through time and known to the researcher and the filer. This refund is to be received in a fixed number of periods, k , after filing. The state variables known to the researcher are $x_{it} = (f_{it}, b_i)$, which is Markov.

We assume that costs of filing have both a fixed and an idiosyncratic component. The fixed costs of filing are denoted by c and the time-varying idiosyncratic shocks are denoted by ε_{it} . These shocks are contemporaneously observed by the filer but unobserved to the researcher. These shocks may depend on the choice of filing and hence we write $\varepsilon(a_{it})$. We assume $\varepsilon(a_{it})$ is independent and identically distributed over time with pdf $g(\varepsilon(a_{it}))$. This is an unknown state variable.

Filer utility is additively separable. The utility of filing in period t is

$$\delta^k b_i - c + \varepsilon(1)$$

when $a_{it} = 1$ and $f_{it} = 1$. The variable δ^k is a k period *exponential* discount factor homogeneous in the population of filers. Utility is $\varepsilon(0)$ if $a_{it} = 0$ and $f_{it} = 1$. As such, the flow utility can be written

$$u(x_{it}, a_{it}) + \varepsilon_{it}(a_{it}) = a_{it} f_{it} (\delta^k b_i - c) + \varepsilon_{it}(a_{it}).$$

With these flow utilities, the filer maximizes the present discounted value of filing-related utilities by choosing α_i^* , a set of decision rules for all possible realizations of observed and unobserved state variables in each time period. That is,

$$\alpha_i^* = \arg \max_{\alpha_i} E_{\alpha_i} \sum_{t=1}^T \delta^{t-1} [u(x_{it}, a_{it}) + \varepsilon_{it}(a_{it})].$$

In Appendix B.2, we implement the methodology of Hotz and Miller (1993); Arcidiciano and Ellickson (2011) with $\varepsilon_{it}(a_{it})$ drawn from the Type-1 extreme value distribution to construct the likelihood function for and exponential discounter's filing in a given period, $a_{it} = 1_{it}$, conditional

on not having filed yet $f_{it} = 1$ as

$$p_{\delta}(a_{it} = 1 | f_{it} = 1, b_i) = \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_{\delta}(a_{i,t+1} = 1 | f_{i,t+1} = 1, b_i))]}.$$

The decision maker trades off the discounted consideration of receiving their refund one day earlier against the discounted considerations of paying their filing costs a day earlier and their forecast that they actually will file the next period, $p_{\delta}(a_{i,t+1} = 1 | f_{i,t+1} = 1, b_i)$. One important observation from Hotz and Miller (1993) is that under rational expectations the one period ahead conditional choice probability can simply be calculated from the data. Given the simplicity of our problem — finite time, simple state space and transition probabilities — one can also simply construct these beliefs at a given parameter constellations, (δ, c) in each iteration of the estimation routine.

Note that we only need to consider those periods up until the time when the person files. Once they file, the utility consequences of filing are eliminated and the likelihood contribution is zero for such observations. Let D_i be the filing date of a given individual. The grand log likelihood is written as

$$\mathcal{L} = \sum_{i=1}^N \left[\sum_{t=1}^{D_i} \ln[p_{\delta}(a_{it} = 1 | f_{it} = 1, b_i)] \right]. \quad (2.2)$$

Identification

From the likelihood of (1), we wish to estimate the parameters δ and c . It is worth noting that exercises of this form suffer generically from identification problems (Rust, 1994). The underlying issue is that monotonic transformations of the instantaneous utility function will yield identical choice rules and hence identical likelihoods. Normalizations and functional form assumptions are often invoked to deliver credible estimates of remaining parameters. One frequent normalization is to assume a specific one period discount factor (Bajari, Benkard and

Levin, 2007; Kennan and Walker, 2011; Aguirregabiria and Mira, 2007).³³

Our environment differs in a compelling way from most exercises in that filing costs and benefits are not experienced in the same period. The fact that $k > 0$ delivers the opportunity to estimate both δ and c from the data. Naturally, this is in the presence of our additional functional form assumptions and normalizing post-filing payoffs to zero.

To understand the intuition of how $k > 0$ allows us to estimate both δ and c lets consider the filling probabilities in time periods $T - 2$ and $T - 1$. Recall that the likelihood for filing in any given period is

$$p_t = \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b - (\delta - 1)c - \delta \ln(p_{t+1})]}, \quad (2.3)$$

where p_t, p_{t+1} are the probabilities of filing in period t and $t + 1$ conditional on not having filed beforehand. In the final period, T , all individuals file, $p_T = 1; \ln(p_T) = 0$, and so the penultimate period filing,

$$p_{T-1} = \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b - (\delta - 1)c]},$$

is driven only by the core incentives of discounted costs and benefits, $[(\delta^{k+1} - \delta^k)b - (\delta - 1)c]$. Two individuals with parameter combinations $(\delta, c) \neq (\tilde{\delta}, \tilde{c})$ such that $[(\delta^{k+1} - \delta^k)b - (\delta - 1)c] = [(\tilde{\delta}^{k+1} - \tilde{\delta}^k)b - (\tilde{\delta} - 1)\tilde{c}] = Q$, will have the same penultimate period filing probabilities. However, moving back one more period reveals the intuition behind the identification. Though these two individuals have the same penultimate period probability, from the perspective of $T - 2$ they assess this future probability differently,

³³The identification of the discount factor in dynamic discrete choice settings has received substantial theoretical attention. Several potential paths forward have been proposed that are not reliant of functional form for identification. One is using variation that changes transition probabilities but do not change contemporaneous utilities (Magnac and Thesmar, 2002). This technique has been usefully applied by Fang and Wang (2015) and Mahajan and Tarozzi (2011) not just for identifying the discount factor, but also for present bias. We are not aware of a variable that can serve such a purpose in our setting.

$$p_{T-2}(\delta, c) = \frac{1}{1 + \exp[Q - \delta \ln(p_{T-1})]} \neq \frac{1}{1 + \exp[Q - \tilde{\delta} \ln(p_{T-1})]} = p_{T-2}(\tilde{\delta}, \tilde{c}).$$

In particular, if $1 > \delta > \tilde{\delta}$ then $p_{T-2}(\delta, c) < p_{T-2}(\tilde{\delta}, \tilde{c})$. That is, the more patient individual appreciates the likelihood of future filing more and is less likely to file at period $T - 2$.

More generally speaking, how do different parameter combinations of (δ, c) effect the evolution of p_{T-t} through each period $T - t$? We know from the above that p_{T-1} is determined by the core incentives of discounted costs and benefits, $[(\delta^{k+1} - \delta^k)b - (\delta - 1)c]$. Furthermore, in Appendix B.3.1 we show that $p_L = \lim_{t \rightarrow \infty} p_{T-t} > 0$ and that

$$(1 - p_L)p_L^{\delta-1} = p_{T-1}^{-1} - 1.$$

That is, while core incentives determine the penultimate period filing probability and the option value it embodies, the weighting of this option value by δ determines the limit. Additionally, in Appendix B.3.2 we demonstrate that distinct pairs of discounting and cost parameters can produce identical filing probabilities on at most one time prior to day T .³⁴ Therefore, we know that for $(\delta, c) \neq (\tilde{\delta}, \tilde{c})$ such that $p_{T-1} = \tilde{p}_{T-1}$ if $\delta > \tilde{\delta}$, then $p_L < \tilde{p}_L$ and $p_{T-t} < \tilde{p}_{T-t}$ for all $t > 1$.³⁵

2.3.2 Estimating Quasi-Hyperbolic Discounting

In our setting, the methodology of dynamic discrete choice can be augmented to estimate present-biased preferences. We assume quasi-hyperbolic $\beta - \delta$ discounting of the form proposed

³⁴In Appendix B.3.3, we also detail this point via simulation. We show in (δ, c) space that level sets of conditional filing probabilities for multiple periods tightly overlap when $k = 0$, indicating that many parameter constellations lead to the same probabilistic choice behavior. When $k > 0$, these level sets separate, implying differential intertemporal patterns of probabilistic behavior for different parameter combinations. Of additional note is that for different values of $k > 0$, the same parameters lead to (at times notably) different filing probabilities. Hence, differential direct deposit use, which generates differences in k , delivers additional identifying variation.

³⁵This follows from equation (2) and the fact that if $p_{T-t} < \tilde{p}_{T-t}$ and $\delta > \tilde{\delta}$, then $-\delta \ln(p_{T-t}) > -\tilde{\delta} \ln(\tilde{p}_{T-t})$.

by Laibson (1997); O’Donoghue and Rabin (1999). An individual is assumed to discount between the present and a future period with discount factor $\beta\delta$, but discount between any two future periods with discount factor, δ alone. If $\beta < 1$ the individual is ‘present-biased’, while if $\beta = 1$, the individual behaves as an exponential discounter.³⁶ The quasi-hyperbolic model elegantly delivers deviations from exponential discounting such as self-control problems, procrastination being one potential manifestation. Because such models feature inconsistencies between long run plans and short run behavior, they require the researcher to provide a formulation of the individual’s beliefs about their own predilection to be present-biased in the future. For our analysis we make the analytically tractable, but admittedly extreme, assumption that individuals are naive with respect to their present bias. That is, they believe that in the future they will behave as if $\beta = 1$. This assumption is used and the simplification it generates for analysis of consumption-savings decisions is discussed in Laibson et al. (2017) and DellaVigna et al. (Forthcoming).³⁷ Our naive present-biased likelihood formulation extends these prior efforts to dynamic discrete choice, providing a purposefully simple method for estimating present bias in such settings.

Incorporating present bias into our estimation procedure requires several steps. First, as noted above, our exponential formulation assumes rational expectations. A naive present-biased individual deviates from rational expectations in the sense that he believes his filing behavior will follow the path of an exponential discounter rather than the true path. As such the true, rational expectations filing probability $p(1_{it}|1_{it}, b_i)$ does not reflect this decisionmaker’s belief. However, the assumption of naivete proves useful in this environment. For a given δ and c , one can solve for the path of beliefs via backwards induction as

$$p_n(a_{it} = 1|f_{it} = 1, b_i) = \frac{1}{1 + \exp \left[(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_n(a_{i,t+1} = 1|f_{i,t+1} = 1, b_i)) \right]}$$

³⁶Additionally, the case of $\beta > 1$ is termed ‘future-biased’.

³⁷It should be noted, however, the assumption of full naivete would be inconsistent with the literature on commitment demand described in footnote 2.

and $p_n(a_{i,T} = 1|f_{i,T} = 1, b_i) = 1$. With $p_n(a_{i,t+1} = 1|f_{i,t+1} = 1, b_i)$ as the forecasted future behavior, we can then estimate β via maximum likelihood under the assumed value of δ and c using similar methods as 2.3.1.³⁸ The likelihood contribution for the critical case of $f_{it} = 1$ becomes

$$p_{\beta\delta}(a_{it} = 1|f_{it} = 1, b_i) = \frac{1}{1 + \exp[\beta(\delta^{k+1} - \delta^k)b_i - (\beta\delta - 1)c - \beta\delta \ln(p_n(a_{i,t+1} = 1|f_{i,t+1} = 1, b_i))]}.$$

In effect, the estimator replaces the rational expectations beliefs inherent to our exponential construction with naive beliefs, and then estimates present bias using these systematically miscalibrated forecasts.

2.3.3 Constructing Counterfactuals

Maximum likelihood estimation of the above dynamic discrete logits provide parameter estimates for exponential discounting and the costs of filing, $\hat{\delta}$ and \hat{c} , and, in the case of quasi-hyperbolic discounting a present bias parameter, $\hat{\beta}$. With these parameters in hand one can construct the counterfactual distribution of filing behavior through backwards induction at the estimated values. For example, under exponential discounting one solves for

$$\hat{p}_{\delta}(a_{it} = 1|f_{it} = 1, b_i) = \frac{1}{1 + \exp[(\hat{\delta}^{k+1} - \hat{\delta}^k)b_i - (\hat{\delta} - 1)\hat{c} - \hat{\delta} \ln(\hat{p}_{\delta}(a_{i,t+1} = 1|f_{i,t+1} = 1, b_i))]} ,$$

³⁸The derivation of this expression is presented in Appendix B.4

with $\hat{p}_\delta(a_{iT} = 1|f_{iT} = 1, b_i) = 1$. The first natural counterfactual distributions for behavior are the average fitted conditional filing probability,

$$\hat{p}_\delta(a_t = 1|f_t = 1) = \frac{1}{N} \sum_{i=1}^N \hat{p}_\delta(a_{it} = 1|f_{it} = 1, b_i)$$

$$\hat{p}_{\beta\delta}(a_t = 1|f_t = 1) = \frac{1}{N} \sum_{i=1}^N \hat{p}_{\beta\delta}(a_{it} = 1|f_{it} = 1, b_i).$$

These counterfactual distributions could be compared to their empirical analog, $\tilde{p}(a_t = 1|f_t = 1)$, for all periods until period $T - 1$.³⁹

From the conditional filing probabilities, one can construct additional counterfactuals for other aspects of filing behavior. First, one can construct the probability of filing unconditional on the contemporaneous filing status, f_{it} , as

$$\hat{q}_\delta(a_{it} = 1|b_i) = \hat{p}_\delta(a_{it} = 1|f_{it} = 1, b_i) \prod_{s=0}^{t-1} (1 - \hat{p}_\delta(a_{is} = 1|f_{is} = 1, b_i)),$$

$$\hat{q}_{\beta\delta}(a_{it} = 1|b_i) = \hat{p}_{\beta\delta}(a_{it} = 1|f_{it} = 1, b_i) \prod_{s=0}^{t-1} (1 - \hat{p}_{\beta\delta}(a_{is} = 1|f_{is} = 1, b_i)).$$

Then one can construct the number and percentage of filers at t (e.g., $\sum_{i=1}^N \hat{q}_\delta(a_{it} = 1|b_i)$ and $\frac{100}{N} \sum_{i=1}^N \hat{q}_\delta(a_{it} = 1|b_i)$ under exponential discounting). Additional useful constructs are the expected average refund value at t ,

$$\hat{b}_{\delta,t} = \frac{\sum_{i=1}^N \hat{q}_\delta(a_{it} = 1|b_i) \times b_i}{\sum_{i=1}^N \hat{q}_\delta(a_{it} = 1|b_i)},$$

$$\hat{b}_{\beta\delta,t} = \frac{\sum_{i=1}^N \hat{q}_{\beta\delta}(a_{it} = 1|b_i) \times b_i}{\sum_{i=1}^N \hat{q}_{\beta\delta}(a_{it} = 1|b_i)},$$

³⁹In period T , the filing probability is assumed to be 1 and all remaining individuals will file by construction.

and the expected filing date for individual i ,

$$\hat{t}_{\delta,i}^* = \sum_{t=0}^T \hat{q}_{\delta}(a_{it} = 1|b_i) \times t,$$

$$\hat{t}_{\beta\delta,i}^* = \sum_{t=0}^T \hat{q}_{\beta\delta}(a_{it} = 1|b_i) \times t.$$

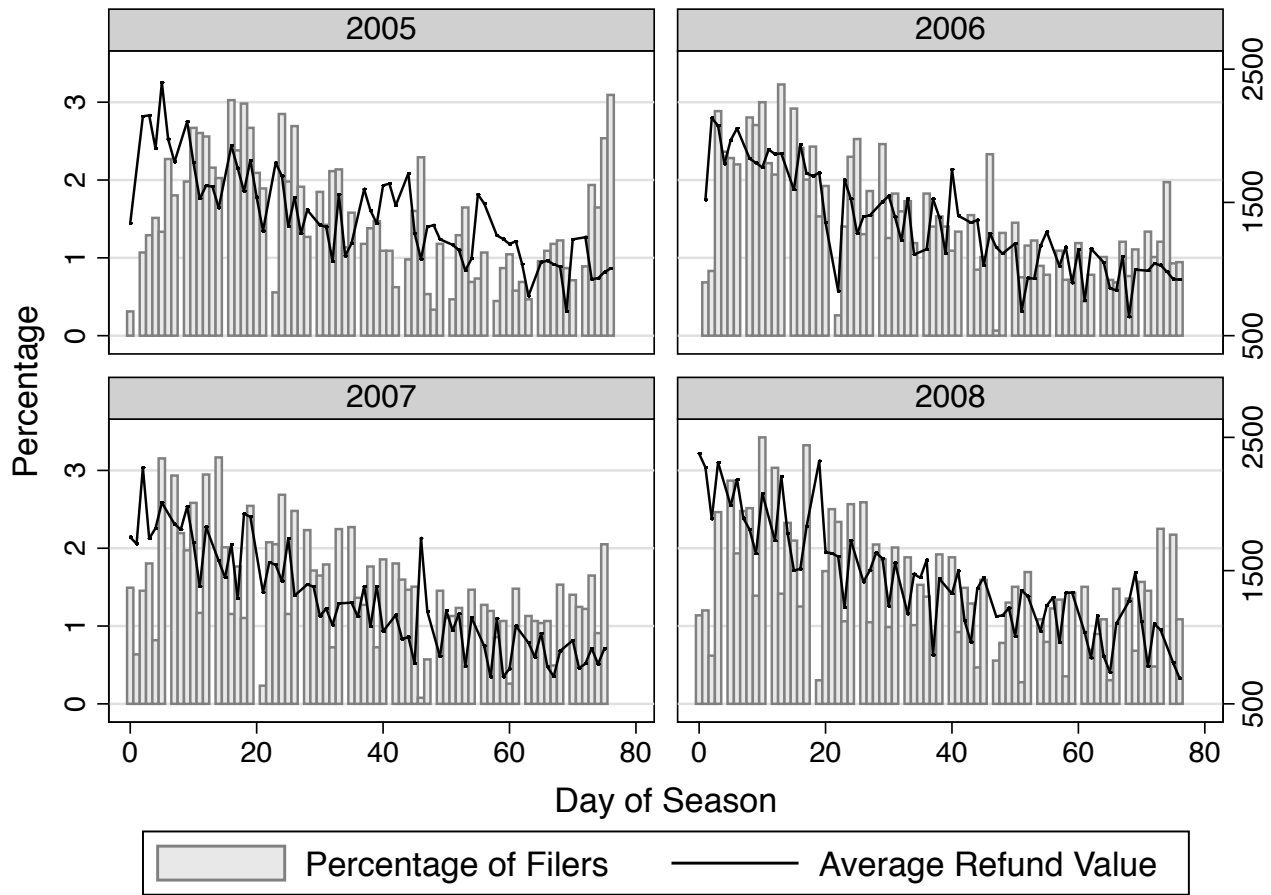
These values, as well, can be compared to their empirical analogs.

2.4 Results

We present the results in four steps: First, we present baseline estimates of exponential discounting and missing mass deviations therefrom. Second, we present estimates of present-biased preferences, rationalizing the deviations from exponential discounting in filing behavior at reasonable parameter values. We also briefly discuss a number of alternative specifications retaining exponential discounting that fail to credibly account for the observed filing patterns. Third, we use the 2008 Stimulus Act to examine the out-of-sample validity of our estimated present-biased model for predicting sensitivities of filing times to changes in filing incentives. Fourth, in a sub-sample of around 1100 subjects we link deviations from exponential discounting in filing behavior to experimental measures of present bias to provide further validation for our behavioral interpretation of the data.

2.4.1 Exponential (δ) Discounting, Quasi-Hyperbolic (β - δ) Discounting, and Excess Mass

Figure 2.1 presents histograms of filing behavior in each year from 2005 to 2008. The figure shows that a large proportion of individuals file early in the tax filing season, with the numbers declining until approximately day 50. From day 50 to the end of the season a pronounced increase in filing is observed. Roughly 9% of filers arrive at VITA sites in the last seven days of



Notes: 2005-2008 percentage of filers on each day of tax season (gray bars) and average refund value for filers on each day (black line).

Figure 2.1: Filing Times and Refund Values

operation each year. Also presented in Figure 2.1 is the average refund value among filers each day. The average value of refunds declines regularly from the beginning of the tax filing season to the end.

Our estimation techniques link refund values, filing times, and the timing of refund receipts via structural models of dynamic discrete choice. A critical component of this procedure is the forecasted future conditional filing probabilities arrived at via rational expectations. In the first approach, we construct bin estimates for conditional filing probabilities in each year

using deciles of refund values.⁴⁰ For each year, on each day VITA sites are open, in each decile of refund value, we calculate the empirical proportion of individuals who have yet to file who file on that day. A total of 4589 bins are constructed delivering corresponding bin estimates for conditional filing probabilities, $\tilde{p}(1_{i,t+1}|1_{i,t+1}, b_i)$. In the second approach, we calculate the stream of forecasts, and hence $\hat{p}(1_{i,t+1}|1_{i,t+1}, b_i)$, in each iteration of the likelihood maximization.

Table 2.2: Aggregate Parameter Estimates 2005 — 2007

	(1)	(2)	(3)	(4)
$\hat{\beta}$	1	1	0.857	0.865
	-	-	(0.011)	(0.011)
$\hat{\delta}$	0.536	0.530	0.99999	0.999993
	(0.003)	(0.007)	-	(0.040)
\hat{c}	3.341	3.40	3.459	3.025
	(0.017)	(0.001)	(0.062)	(0.082)
# Observations	1,010,387	1,038,245	1,038,245	1,038,245
Log-Likelihood	-65402.38	-67565.92	-65950.99	-65918.86

Notes: Column 1 provides maximum likelihood estimates for δ and cost of filing with observed filing probabilities ($\beta = 1$ assumed). Column 2 provides estimates for δ and cost with simultaneous forecasts. Column 3 estimates β and cost with simultaneous forecasts (δ fixed). Column 4 estimates β , δ and cost with simultaneous forecasts. δ is estimated as $\delta = \frac{1}{1+e^\phi}$ in columns 2, and 4. In columns 2 and 4 the standard error for ϕ is reported.

Table 2.2, columns (1) and (2) present aggregate parameter values based on these two techniques under the assumption of exponential discounting using the data from 2005 to 2007. The data from 2008 and out-of-sample analysis of the 2008 Stimulus Payments are presented in section 2.4.2. In column (1) we estimate both the discount factor, δ , and filing cost, c ,

⁴⁰One minor hurdle to creating such values is that VITA sites are closed on holidays and Sundays. We address this by altering our measure of time to reflect only those days where the VITA sites are open each year. A total of 65 days in the tax filing season remain. Note that recognizing a change in the effective timing of choices also requires us to change the intertemporal tradeoffs built in to our estimator. Recall from section B.2.1 that we are able to assume that the terminal option is taken in the next period. To account for days that the VITA sites are closed on Sundays and holidays we simply assume that the next period is two periods away resulting in the conditional probability $p(1_{it}|1_{it}, b_i) = \frac{1}{1+\exp[(\delta^{k+2}-\delta^k)b-(\delta^2-1)c-\delta^2\ln(p(1_{i,t+2}|1_{i,t+2}, b_i))]}$.

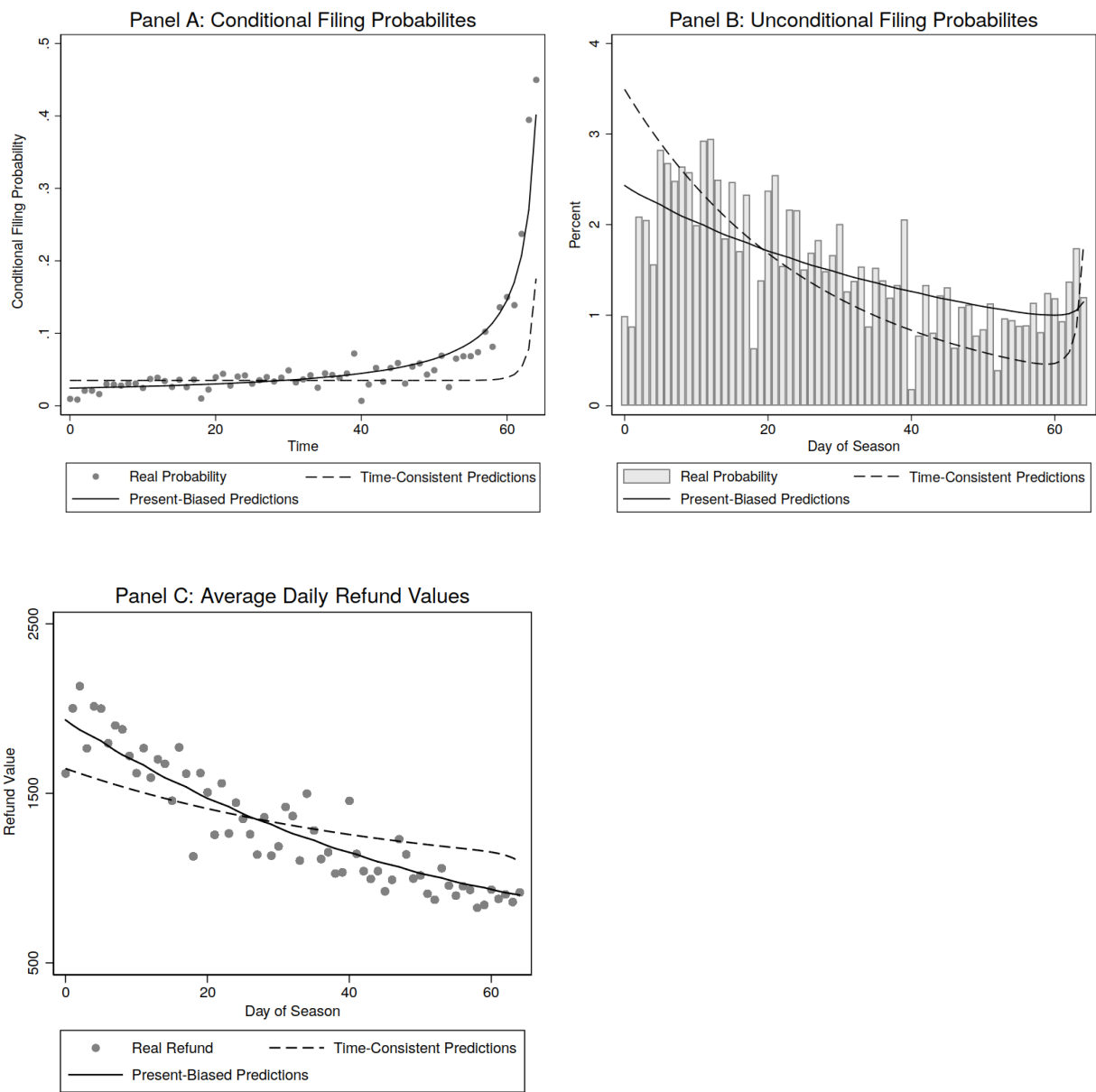
finding an average filing cost of $\hat{c} = 3.341$ and a discount factor of $\hat{\delta} = 0.536$.⁴¹ In column (2), with forecasts $\hat{p}(1_{i,t+1}|1_{i,t+1}, b_i)$ generated in each iteration of the likelihood, we find virtually identical results.

In order to capture empirical regularities of not disproportionately filing early, individuals must substantially discount their filing benefits. Hence, receiving a sizable refund in several weeks' time can be outweighed by modest filing costs. Estimates of discount factors in the range observed from Table 2.2, columns (1) and (2) imply discount rates far below empirical rates of interest. When such extreme impatience is required to rationalize empirical behavior in field settings, researchers often appeal to calibrational arguments to reject exponential discounting (Fang and Silverman, 2009; Shapiro, 2005). In contrast, in experimental settings it is not unusual to identify individual discount rates in the hundreds of percents per year (see, e.g., Frederick, Loewenstein and O'Donoghue, 2002). Our exercise takes the exponential estimates as a valid benchmark and examines the adherence of true intertemporal patterns of filing to predicted behavior under the estimated exponential model.

Figure 2.2 presents predicted and actual filing behavior for 2005-2007. In Panel A, we examine average predicted and actual conditional filing probabilities. The dashed line represents the exponential prediction, $\hat{p}_\delta(a_t = 1|f_t = 1)$, using the estimates from from Table 2.2, column (2). Predicted and actual conditional filing probabilities correspond closely early in the filing season, but diverge as the tax deadline approaches. On the final two days, the true filing probability exceeds $\hat{p}_\delta(a_t = 1|f_t = 1)$ by around 30 percentage points.⁴² The exponential

⁴¹As discussed in section 2.3.1, our discount factor has two sources of identification. The first comes from the formulation of the problem, yielding differences between the timing of costs and benefits. Without $k > 0$, separate estimation of δ and c would be impossible. In Appendix Table B.4, we reconstitute the analysis of Table 2.2 with the assumption of $k = 0$ for all filers. These estimates show the hallmarks of identification problems with flat likelihoods, sensitivity to starting values and, at times, failed convergence. The second comes from the timing impacts of direct deposit. In Appendix Table B.5 and B.6, we provide separate estimates for individuals with and without direct deposit, showing qualitatively similar, though substantially less precise estimates across the two groups.

⁴²It is important to note that the presented counterfactual distribution is an in-sample prediction. Exercises of missing mass generally leave out a region of interest for estimation. As the focus of our project is procrastination, our primary region of interest is the last seven days of the tax-filing season. In Appendix A.1, we reconstitute our estimation removing the last seven days of the tax filing season (Table B.7). As one might expect, misprediction at the end of the tax filing season increases substantially when the last seven days are excluded from estimation. In order to provide more conservative estimates for missing mass we use the full data set for estimation.



Notes: Panel A: 2005-2007 real conditional filing probabilities, $\bar{p}(1_t|1_t)$, predicted time-consistent filing probabilities, $\hat{p}_\delta(1_t|1_t)$, and predicted present-biased filing probabilities, $\hat{p}_{\beta\delta}(1_t|1_t)$, throughout the tax season with $t = 64$ corresponding to the day before the tax deadline. Panel B: 2005-2007 real unconditional filing probabilities, $\bar{q}(1_t|1_t)$, predicted time-consistent unconditional filing probabilities, $\hat{q}_\delta(1_t|1_t)$, and predicted time-inconsistent unconditional filing probabilities, $\hat{q}_{\beta\delta}(1_t|1_t)$. Panel C: Observed average daily refund values and predicted average daily refund for time-consistent and present biased models. All predicted values generated from exponential discounting with $\hat{\delta} = 0.530$ and $\hat{c} = 3.40$ (Table 2.2, column (2)) and from quasi-hyperbolic discounting with $\hat{\beta} = 0.865$, $\hat{\delta} = 0.999993$ and $\hat{c} = 3.025$ (Table 2.2, column (4)).

Figure 2.2: Predicted and Actual Filing Behavior

counterfactual dramatically underpredicts filing probabilities at the end of the tax season.⁴³

Figure 2.2, Panels B and C present daily filing percentages and average refund values for the days of VITA site operation, and also provide corresponding model predictions. Dashed lines in both panels correspond to the exponential predictions $\frac{100}{N} \sum_{i=1}^N \hat{q}_\delta(a_t = 1|b_i)$ and $\hat{b}_{\delta,t}$, respectively. The estimated exponential model over-predicts the percentage of individuals who file early in the season, with substantial under-prediction as the tax deadline approaches. Interestingly, because the model predicts so few people will file during the middle of the tax season, it provides a slight overestimate for the number of filers the day before the deadline. Panel C also highlights that the aggregate estimates predict less sensitivity of filing times to refund values than exists in the data. This result is sensible: given the high degree of impatience, dissimilar refund values have quite similar discounted implications.⁴⁴

Table 2.3, Panel A presents estimates of excess mass relative to the exponential benchmark for both conditional and unconditional filing probabilities.⁴⁵ On average, from $T - 7$ to $T - 1$, $\tilde{p}(a_t = 1|f_t = 1)$ exceeds $\hat{p}_\delta(a_t = 1|f_t = 1)$ by nearly 16%-age points. Implementing the bootstrap procedure of Efron and Tibshirani (1994), we find the difference between observed and predicted behavior to be highly significant. Stated in unconditional terms, from $T - 7$ to $T - 1$ an average of 0.65% of filers are predicted to arrive each day, while 1.18% actually do. This equates to around 82% more filers than expected by our counterfactual exponential predictions.

⁴³In Appendix Figure B.2, we present an alternate counterfactual, $\hat{p}(1_t|1_t)$, based on the empirical one-period ahead conditional filing probabilities as opposed to backwards induction. Such a counterfactual makes use of the true one-period ahead behavior, rather than the model's prediction and hence delivers a much less smooth counterfactual distribution. Relative to this benchmark, as well, substantial missing mass is observed.

⁴⁴In Appendix Figures B.3 and B.4, we reconstruct Figure 2.2 separately for individuals who receive their refund by paper check versus direct deposit. This figure demonstrates that the 60% of individuals receiving paper check are predicted to have virtually constant refund values throughout the tax season. For such a high degree of impatience almost all values of refunds are discounted to a common base of zero. The model fit for direct deposit refund recipients is substantially better.

⁴⁵To construct these tests we use a standard stratified bootstrap (Efron and Tibshirani, 1994). Our bootstrap proceeds in four stages. First, stratifying by year, we independently resample the data set 500 times, constructing each time the one period ahead conditional filing probability bin estimates for the given sample. Second, we implement our maximum likelihood estimation on each sample. This yields 500 estimates of $\hat{\delta}$ and \hat{c} . Third, using these estimates we generate 500 counterfactual distributions for filing behavior. Fourth, we take the difference between the resampled observed filing behavior and counterfactual filing behavior. This yields a distribution of 500 missing mass estimates upon which statistical tests can be implemented.

Table 2.3: Excess Mass Results

	Excess Conditional Probability $\tilde{p}(a_t = 1 f_t = 1) - \hat{p}(a_t = 1 f_t = 1)$	Excess Unconditional Probability $\tilde{q}(a_t = 1 f_t = 1) - \hat{q}(a_t = 1 f_t = 1)$
<i>Panel A: Exponential Predictions:</i>		
$T - 1 : T - 7$ Average	0.157*** (0.004)	0.0053*** (0.0002)
<i>Panel B: Quasi-Hyperbolic Predictions:</i>		
$T - 1 : T - 7$ Average	0.0218*** (0.002)	0.0018*** (0.0004)

Notes: Estimates of excess conditional filing probability, $\tilde{p}(1_t|1_t) - \hat{p}(1_t|1_t)$, and the excess unconditional filing probability, $\tilde{q}(1_t|1_t) - \hat{q}(1_t|1_t)$, over the seven days prior to the tax deadline. Bootstrapped standard errors in parentheses from 500 bootstrap samples.

The data compellingly reject the estimated model of exponential discounting with $\hat{\delta} = 0.530$ and $\hat{c} = 3.40$.⁴⁶

The results of Figure 2.2 and Table 2.3 demonstrate substantial deviations between actual behavior and predictions based on exponential discounting. Table 2.2, columns (3) and (4) present estimates of quasi-hyperbolic discounting using the method described in section 2.3.2. In column (3) we fix $\delta = 0.99999$ and estimate present bias, β and costs, c . Costs are estimated to be close to the prior estimates, $\hat{c} = 3.46$, and present bias is estimated to be $\hat{\beta} = 0.86$. Interestingly, this value of present bias is close to recent empirical estimates for present bias from intertemporal effort choices (Augenblick, Niederle and Sprenger, 2015; Augenblick and Rabin, 2016).⁴⁷ In column (4), we conduct the same analysis estimating δ within the likelihood, and reach quite similar conclusions.⁴⁸

Allowing for present bias dramatically alters the extent of excess mass at the end of the tax filing season. In Figure 2.2, we conduct predictions for conditional filing, unconditional filing,

⁴⁶Figure 6, Panel A presents the inconsistency between exponentially predicted and actual values in a complementary way. Using predicted and actual filing days, it demonstrates that from 2005 to 2007 the estimated values, $\hat{\delta} = 0.530$ and $\hat{c} = 3.40$, lead to predicted filing times that are 2 to 4 days earlier, on average, than actual filing times.

⁴⁷Augenblick, Niederle and Sprenger (2015) estimate $\beta = 0.88$ and Augenblick and Rabin (2016) estimate $\beta = 0.83$ from intertemporal effort choices.

⁴⁸Explanation of challenges for starting values.

and estimated refund values over time for the estimates of Table 2.2 , column (3). The solid lines in each panel represent these estimates. Quasi-hyperbolic discounting with a relatively modest degree of present bias, $\beta = 0.86$, matches intertemporal patterns in conditional and unconditional filing probabilities as well as the evolution of refund values through time.⁴⁹ Table 2.3 contrasts the excess mass identified under exponential discounting and quasi-hyperbolic discounting. The average excess conditional filing probability in the final week reduces by around 85% under the quasi-hyperbolic formulation, while the excess unconditional filing probability reduces by around two-thirds. Reasonable parameter values of present bias, close to other recent estimates, provide a credible account for empirical filing behavior.

The results to here indicate that our formulation of exponential discounting fails to match key intertemporal patterns in filing behavior, while quasi-hyperbolic discounting provides a much more coherent account of the data. Because the quasi-hyperbolic discounting model appeals to an additional degree of freedom, its improved fit should be expected. A natural question is whether alternate degrees of freedom, retaining the exponential formulation, can also provide a high degree of in-sample fit. In Appendix B.1, we evaluate a number of exercises investigating such degrees of freedom in the form of extreme costs, extreme shocks, and heterogeneity in discounting and costs along observables. In each case, key patterns of filing behavior remain unaccounted for. Nonetheless, it must be recognized that improved in-sample fit for the behavioral model relative to the neoclassical benchmark proves only the value of that specific degree of freedom and does not constitute a test of the model. In the next sections, we move out-of-sample to provide the required tests.

⁴⁹In Appendix Figures B.6 and B.7 we reconstruct Figure 2.2 separately for individuals who receive their refund by paper check versus direct deposit. This figure demonstrates that the behavior of both paper check and direct deposit refund recipients is well matched by the aggregate estimates, with sensitivity in filing times to refund values predicted for both groups.

2.4.2 Response to Changing Incentives: 2008 Stimulus Payments

Rationalizing behavior in-sample with $\beta - \delta$ discounting is important, but does not constitute a stringent test of present-biased preferences. Given that the behavioral model adds a parameter to the standard exponential formulation, one should not be overly surprised by its improved fit. Here, we examine responses to the exogenous changes in filing incentives imposed by the Economic Stimulus Act of 2008 to provide an out-of-sample validation for our behavioral interpretation.

In Figure 2.3, we provide a difference-in-difference investigation of whether receiving a 2008 Stimulus Payment induced earlier filing. Using the parameters of the 2008 Stimulus Act, we calculate (without adjusting for inflation) the value of each individual's potential Stimulus Payment for 2005, 2006, and 2007. We present mean filing dates (solid lines in both panels) for individuals who would and would not receive Stimulus Payments from 2005 to 2008. In years prior to 2008, those with and without potential Stimulus Payments follow very similar trends with potential recipients filing earlier.⁵⁰ In 2008, the trends diverge. Those without Stimulus Payments in 2008 continue to file later in the tax season, while those with Stimulus Payments file earlier.

Given the apparent response to the 2008 Stimulus Payments, a natural question is how well 2008 filing behavior is predicted both in general and with respect to the sensitivity of filing to Stimulus Payments. For both our estimated benchmark model of exponential discounting, Table 2.2, column (1), and present bias, Table 2.2, column (3), we can predict each individual's conditional filing probability in 2008, $\hat{p}_\delta(1_{it}|1_{it}, b_i)$ and $\hat{p}_{\beta\delta}(1_{it}|1_{it}, b_i)$, respectively. Similarly, we can construct expected filing days, $\hat{t}_{\delta,i}^*$ and $\hat{t}_{\beta\delta,i}^*$.⁵¹ Hence, we can examine whether our estimates reliably reproduce the patterns in filing behavior in general and the apparent sensitivity to Stimulus Payments.

⁵⁰The earlier filing of potential recipients is likely due to the Stimulus Payment's income thresholds and the larger proportional refunds for lower-income individuals in the sample.

⁵¹These values take into account for each individual not only their standard refund, but also their projected 2008 Stimulus Payment. In order to develop this projection, we assume that the Stimulus Payment will be received 14

Along with true filing times, in Figure 2.3 we present (dotted lines) mean predicted filing times under exponential and $\beta - \delta$ discounting, respectively. These predictions are in-sample for 2005-2007 and out-of-sample for 2008. Echoing our missing mass analysis, Panel A demonstrates that the exponential model's predicted filing times are two to four days earlier than true filing times during the in-sample period. Importantly, because of the extreme estimated degree of impatience under exponential discounting, the model predicts effectively no responsiveness to the changing incentives generated by the 2008 Stimulus Payments. In contrast, the estimated model of present bias both matches the difference between potential recipients and non-recipients from 2005-2007 and closely predicts the observed sensitivity to stimulus receipt.

Table 2.4 shows the results of Figure 2.3 in more detail. Column (1) provides corresponding regression analysis of the true day of filing, t_i^* . The results show that receiving a Stimulus Payment in 2008 (coefficient of variable *Interaction*) induced tax filers to arrive 1.92 (s.e. = 0.98) days earlier.⁵² Adding control variables in columns (2) and (3) does not change the results dramatically. Potential stimulus recipients file earlier in the years 2005 to 2007 and disproportionately so in 2008, the year of the stimulus.

Table 2.4, columns (4)-(9) provide predicted difference-in-difference estimates for the exponential and $\beta - \delta$ predictions – corresponding to the dotted lines in Figure 2.3. In contrast to the exponential predictions of columns (4)-(6), columns (7)-(9) demonstrate that $\beta - \delta$ discounting delivers a plausible account both of the average filing times and of the sensitivity of filing times to the 2008 Stimulus Payments. The predicted mean response to the stimulus of 2.34 (s.e. = 0.10) days in column (7) under $\beta - \delta$ preferences closely matches the true sensitivity of around 2 days. Appendix Figures B.8 and B.9 reproduce the Figures 2 and 5 for days after the refund. For example for exponential discounting,

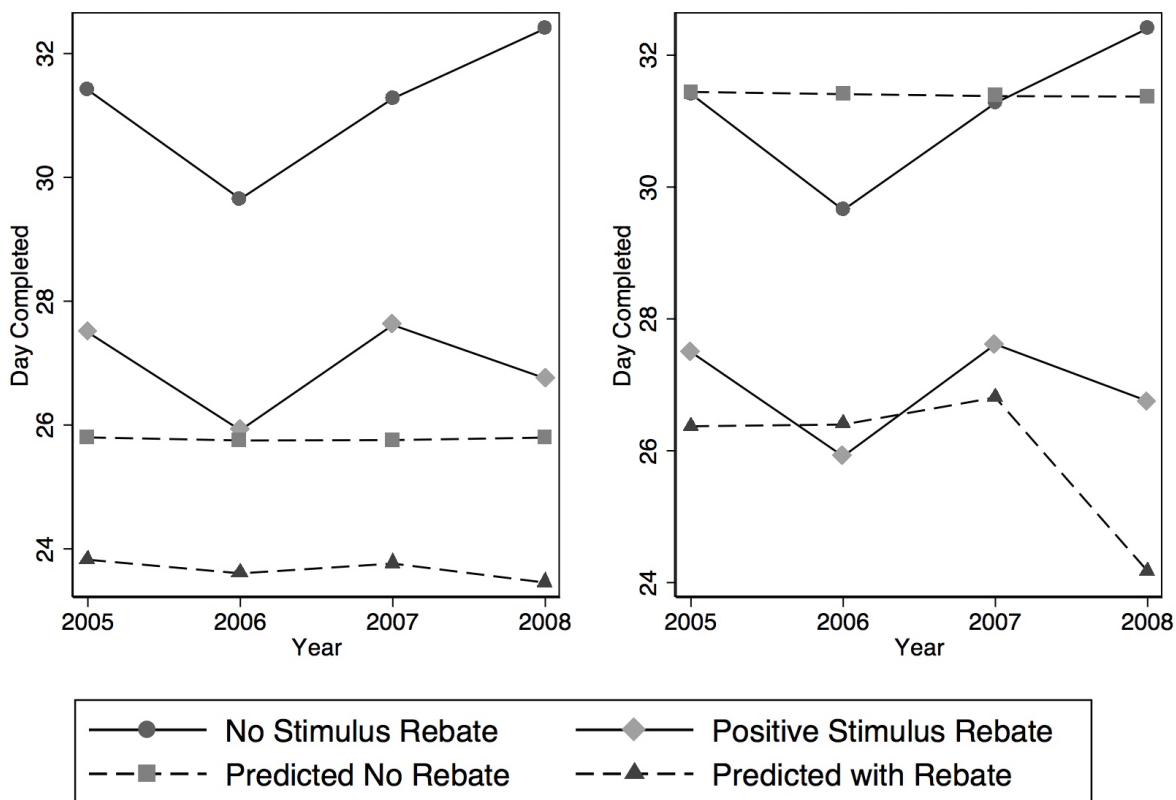
$$\hat{p}(1_{it}|1_{it}, b_i) = \frac{1}{1 + \exp \left[(\hat{\delta}^{k+1} - \hat{\delta}^k)b + (\hat{\delta}^{k+14+1} - \hat{\delta}^{k+14})r - (\hat{\delta} - 1)\hat{c} - \hat{\delta} \ln(\hat{p}(1_{i,t+1}|1_{i,t+1}, b_i)) \right]},$$

where r is the Stimulus Payment, with $\hat{p}(1_{iT}|1_{iT}, b_i) = 1$.

⁵²Appendix Table B.8 provides a placebo test, turning on the Stimulus Payments in 2007 as opposed to 2008. Null effects of Stimulus Payments are observed, supporting the view that earlier filing dates for payment recipients in 2008 are truly due to the Stimulus Payment and not other factors.

Panel A: Exponential Prediction (\hat{t}^*)

Panel B: Present-Biased Prediction ($\hat{t}_{\beta\delta}^*$)



Notes: Mean filing dates, t^* , for potential stimulus recipients (diamonds) and non-recipients (circles) from 2005 to 2008. Panel A: predicted filing date, \hat{t}^* , from exponential discounting for stimulus recipients and non-recipients with $\hat{\delta} = 0.536$ and $\hat{c} = 3.341$, from Table 2.2, column (1). Predictions are in-sample for 2005-2007 and out-of-sample for 2008, the year of the Stimulus Payments. Panel B: predicted filing date, $\hat{t}_{\beta\delta}^*$, from quasi-hyperbolic discounting with $\delta = 0.99999$, $c = 3$ and $\beta = 0.92$, from Table 2.2, column (4). Predictions are in-sample for 2005-2007 and out-of-sample for 2008, the year of the Stimulus Payments.

Figure 2.3: Filing Dates, Stimulus Payments, and Out-of-Sample Predictions

filing probabilities and refund values in 2008 under the two models, again showing improved out-of-sample prediction for present bias.

Table 2.4: Predicted and Actual Difference-in-Difference Effect of Stimulus Payments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Dep. Var.</i>	t_i^* Observed Filing Day			\hat{t}_i^* Exponential Predicted Filing Day			$\hat{t}_{i,\beta\delta}^*$ $\beta-\delta$ Predicted Filing Day		
1Stimulus Payment > 0	-3.731*** (0.545)	-1.079* (0.574)	-0.797 (0.593)	-2.044*** (0.042)	0.690*** (0.044)	0.782*** (0.046)	-4.852*** (0.064)	-0.229*** (0.020)	-0.155*** (0.017)
1Year = 2008	1.682* (0.944)	1.781* (0.945)	1.744* (0.943)	0.031 (0.022)	0.119 (0.075)	0.114 (0.074)	-0.036 (0.074)	0.056 (0.041)	0.048 (0.034)
Interaction	-1.919* (0.983)	-1.647* (0.982)	-1.766* (0.981)	-0.299*** (0.056)	-0.073 (0.080)	-0.062 (0.079)	-2.346*** (0.103)	-2.136*** (0.044)	-2.154*** (0.037)
Constant	30.722*** (0.520)	31.619*** (0.523)	31.618*** (0.527)	25.766*** (0.015)	26.664*** (0.036)	26.691*** (0.036)	31.408*** (0.029)	32.009*** (0.005)	32.008*** (0.008)
# Observations	22525	22525	22525	22525	22525	22525	22525	22525	22525
Control Set 1	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Control Set 2	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Ordinary least squares regression for difference-in-difference effect of Stimulus Payment. Columns (1-3): Dependent variable is actual filing day, t_i^* . Columns (4-6): Dependent variable is predicted filing day, \hat{t}_i^* under exponential discounting with $\hat{\delta} = 0.536$ and $\hat{c} = 3.341$, from Table 2.2, column (1). Columns (7-9): Dependent variable is predicted filing day, $\hat{t}_{i,\beta\delta}^*$, from quasi-hyperbolic discounting with $\delta = 0.99999$, $c = 3$ and $\beta = 0.92$, from Table 2.2, column (4). Columns (4-6) and (7-9) predictions are in-sample for 2005-2007 and out-of-sample for 2008, the year of the Stimulus Payments. Standard errors in parentheses.

Control Set 1: Refund, AGI, Direct Deposit Status.

Control Set 2: Number of Dependents, Taxable Income, Earned Income, Social Security Benefits, Filing Status Binaries (2008 stimulus act determinants).

Levels of significance: *, **, *** for 10%, 5%, and 1%, respectively.

2.4.3 Experimental Present Bias and Tax Filing Behavior

In addition to examining out-of-sample predictive power for $\beta - \delta$ in filing behavior, our environment provides independent means for identifying present-biased preferences. As described in detail in section 2.2.3, we elicited time preferences experimentally for a subsample of around 1100 individuals in 2007 and 2008.

For each experimental subject we can link their experimental measure of present bias, $Present\ Biased_i$, to their individually fitted pattern of exponential conditional filing probabilities, $\hat{p}(1_{it}|1_{it}, b_i)$, and their exponential expected filing date, \hat{t}_i^* .⁵³ Figure 2.4, presents real and

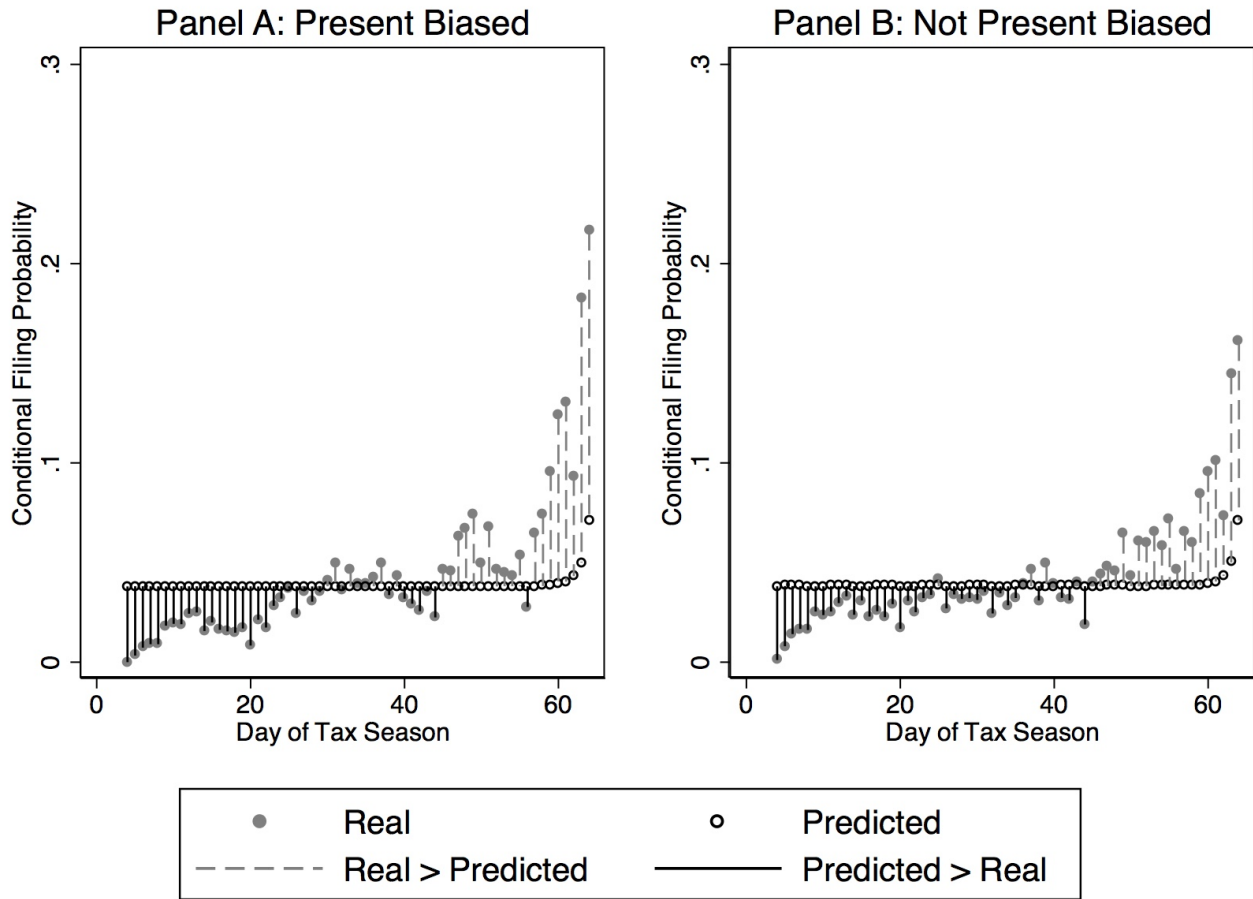
⁵³In order to construct the predicted conditional filing probabilities in-sample for each individual, we re-estimate the baseline exponential model of Table 2.2, column (1) including the 2008 data (and the Stimulus Payments). The corresponding parameter values are effectively unchanged from those reported in Table 2.2, column (1) with $\hat{\delta} = 0.527$ and $\hat{c} = 3.295$. From these estimates we construct the counterfactual $\hat{p}(1_{it}|1_{it}, b_i)$ for each individual via backwards induction.

predicted conditional filing probabilities separately for present-biased and not present-biased subjects. Because our experiment is conducted at only one VITA site, which was not open every day of the tax season, a number of days have no individuals filing. Hence, Figure 2.4 provides 5 day running averages beginning at the tax deadline and moving backwards.⁵⁴

Different patterns of mis-prediction are apparent across present-biased and non present-biased subjects. Though predicted conditional filing probabilities are similar across the two groups, real behavior is markedly different. Present-biased subjects file with higher probability close to the filing deadline and provide greater deviations between predicted and actual behavior than their non present-biased counterparts.

Table 2.5 provides corresponding analysis. In column (1), we regress predicted exponential conditional filing probabilities, $\hat{p}_t(1_{it}|1_{it}, b_i)$, on the day of the tax season, an indicator for *Present Biased*, and their interaction. Present-biased subjects do not differ in terms of their predicted filing dates at any part of the tax season. In column (2), we present an identical regression with true conditional filing probabilities, $\tilde{p}(1_{it}|1_{it}, b_i)$, as dependent variable. Present-biased subjects indeed file with lower probability at the beginning of the tax season and with greater probability at the end of the tax season. In column (3), we present regressions with the difference, $\tilde{p}(1_{it}|1_{it}, b_i) - \hat{p}(1_{it}|1_{it}, b_i)$, as dependent variable. Present-biased subjects are found to file with lower probability than expected early in the tax filing season and with higher probability than expected late in the tax filing season. Columns (4)-(6) repeat this analysis with the exponentially predicted and actual filing times, \hat{t}_i^* and t_i^* , as dependent variables. Experimentally measured present bias is predictive of filing later in the tax season and filing later than the exponential model prediction. Table 2.5 demonstrates the plausibility of interpreting our measures of missing mass as evidence of procrastination. Experimental measures of present bias are tightly linked to the exponential model's mis-prediction of filing behavior close to the tax deadline.

⁵⁴The unsmoothed figure is provided in Appendix A.1 Figure B.10 and presents qualitatively similar, albeit noisier results.



Notes: 2007-2008 predicted and real conditional filing probabilities, $\hat{p}(1_t|1_t)$ and $\bar{p}(1_t|1_t)$, throughout the tax season with $t = 64$ corresponding to the day before the tax deadline. Panel A: Present-Biased Subjects ($n= 360$). Panel B: Not Present-Biased Subjects ($n= 754$) All predicted values generated from exponential discounting with $\hat{\delta} = 0.527$ and $\hat{c} = 3.295$. Data smoothed with 5 day running average.

Figure 2.4: Predicted and Actual Filing Probabilities by Present Bias

Table 2.5: Present Bias and Deviations from Exponential Prediction

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Conditional Filing Probability			Filing Day		
	Predicted $\hat{p}_t(1_{it} 1_{it}, b_i)$	Observed $\bar{p}(1_{it} 1_{it}, b_i)$	Difference $\bar{p}(1_{it} 1_{it}, b_i) - \hat{p}_t(1_{it} 1_{it}, b_i)$	Predicted \hat{t}_i^*	Observed t_i^*	Difference $t_i^* - \hat{t}_i^*$
Present Biased (=1)	0.0002 (0.0005)	-0.0081*** (0.0023)	-0.0083*** (0.0024)	-0.0121 (0.1687)	2.5545** (1.1365)	2.5666** (1.1156)
Present Biased x Day	0.0000 (0.0000)	0.0003** (0.0001)	0.0003** (0.0001)	-	-	-
Day	0.0000 (0.0000)	0.0011*** (0.0001)	0.0011*** (0.0001)	-	-	-
Constant	0.0372*** (0.0004)	0.0072*** (0.0015)	-0.0300*** (0.0015)	23.5255*** (0.0987)	31.1844*** (0.6722)	7.6588*** (0.6556)
R-Squared	0.002	0.012	0.012	0.000	0.004	0.005
# Filers	1114	1114	1114	1114	1114	1114
# Observations	36736	36736	36736	1114	1114	1114

Notes: Ordinary least squares regression for relationship between experimental measure of present bias, predicted and actual filing behavior. Columns (1) through (3): Dependent variables constructed from exponentially predicted and observed conditional filing probabilities, $\hat{p}_t(1_{it}|1_{it}, b_i)$ and $\bar{p}(1_{it}|1_{it}, b_i)$. Columns (4) through (6): Dependent variables constructed from predicted and observed filing times: \hat{t}_i^* and t_i^* . Predicted values constructed under exponential discounting with $\hat{\delta} = 0.527$ and $\hat{c} = 3.295$. Standard errors in parentheses, clustered by individual in columns (1) through (3), robust in columns (4) through (6).

Levels of significance: *, **, *** for 10%, 5%, and 1%, respectively.

2.5 Discussion and Conclusion

Procrastination is a critical deviation between long term plans and short term behavior, inconsistent with the neoclassical formulation of exponential discounting. Recognizing the challenges of identifying behavioral models of such deviations in field data, this paper presents a way to potentially identify present-biased procrastination. Combining methods from structural estimation of dynamic discrete choice and inference from missing mass, we show a way in which deviations from a specific model of exponential discounting can be measured. We compare realized tax-filing behavior to counterfactual estimates from an optimal stopping problem developed under the assumption of exponential discounting. In a sample of around 22,000 low-income tax filers, realized and exponentially-predicted distributions of filing probabilities differ dramatically as the tax deadline approaches. The existence and location of the missing mass in tax filing behavior is consistent with quasi-hyperbolic discounting with a moderate degree of present bias. This interpretation of present bias is bolstered by out-of-sample prediction for the sensitivity

of filing times to the 2008 stimulus payments, and incentivized experimental measures of time preferences.

The results presented here rely on linkages across a diverse set of prior research and may, similarly, have implications across a range of fields. First, the paper shows that using structural estimation to develop counterfactual distributions for behavior may valuably expand the scope of missing mass exercises and sharpen corresponding tests. When using a structural benchmark, one need not rely (as is tradition) on a smooth counterfactual distribution through a point of potentially changing incentives to identify missing mass. Further, one knows precisely the model rejected when realized deviations are observed. Though these points are general, we believe that specific behavioral applications in intertemporal choice could be readily implemented. Stopping problems of the form examined here such as submitting applications, paying bills, changing retirement plans or credit cards, and mortgage refinance abound, and corresponding behaviors are often anecdotally linked to present-biased preferences. A valuable line of research could emerge from precise structural investigation of the extent to which behavior deviates from exponential benchmarks in these settings. While we show that our missing mass is compellingly rationalized by a form of quasi-hyperbolic discounting, many alternative models exist. Future work should investigate other models of time preferences and relax our critical assumption of naïveté.

Second, the paper provides both a rationalization for the observed deviations from exponential discounting and a corresponding out-of-sample validation exercise. Given that many candidate theories can rationalize an observed deviation from a given model, such tests should be viewed as a necessary step to exercises of this form. Different candidate theories could have quite different predictions or welfare implications and so knowing the most appropriate one becomes a critical input both for further study and policy discussion. We believe our exercise shows a path forward particularly for exercises that ascribe behavioral motivations to findings of missing mass. When candidate rationalizations include behavioral models with additional free parameters, better in-sample fit should be viewed as relatively weak evidence of model success.

Third, our interpretation of present bias is bolstered in this paper by incentivized experi-

mental measures. A prominent discussion in the behavioral literature has questioned the use of monetary discounting measures, like ours, for the study of time preferences. In principle, the fungibility of money renders such choices useless for informing researchers about consumption preferences, unless subjects are liquidity constrained. Predicting individual differences in procrastination in the timing of tax-filing for a sample of likely constrained individuals — presumably a consumption preference related to labor-leisure tradeoffs— highlights the possibility of using such methods at times as a measure of preferences. Clearly, more research is needed to fully understand how, when, and why monetary measures of discounting convey valuable information on preferences.

Last, the paper can identify and quantify procrastination in the field for the low income sample in question. Our results show that roughly seventy-one percent of filers in the last week before the deadline are potential procrastinators. Though we cannot extrapolate from this prevalence among low-income filers to the broader population, there may be reasons to be interested in this population specifically. Low income individuals access a number of government services in a similar way to our intertemporal filing problem. Benefits such as food stamps must be signed up for initially and the spending of corresponding receipts is a secondary intertemporal problem. If the level of present bias measured in our sample is indicative of behavior in accessing and spending such benefits, policy makers may well be interested in altering the timing of costs and benefits to achieve coverage and smoothing objectives. Such targeted policy discussions are now equipped with a key observation on the potential extent of procrastination and present bias.

Acknowledgements

Chapter 2, in full, is currently being prepared for submission for publication of the material. Martinez, Seung-Keun; Charles, Sprenger; Stephan, Meier. “Procrastination in the Field: Evidence from Tax-Filing” The dissertation author was one of three primary investigators and authors of this material.

Chapter 3

A Test for Risk-Averse Expected Utility

3.1 Introduction

The recent contribution of Kubler, Selden and Wei (2014) provides a GARP-like test for risk-averse expected utility maximization in a contingent-consumption environment. In an environment with a single consumption good and finite states of the world, they establish an acyclicity condition on observed data which is both necessary and sufficient for a finite list of observed price and consumption pairs to be consistent with the hypothesis of expected utility maximization. Thus, their paper provides a counterpart of the classical work of Afriat (1967) with the added restriction that rationalizations be risk-averse expected utility.

As Kubler, Selden and Wei (2014) note, their test is universal in nature, removing all existential quantification. Their test amounts to verifying that the product of certain cycles of risk-neutral prices be bounded above by one. Our aim in this note is to provide a *different* universal test. Our test should be distinguished from the Kubler, Selden and Wei (2014) test in three ways. First, it applies to any finite number of consumption goods, whereas the test of Kubler, Selden and Wei (2014) only applies for a single consumption good. Secondly, our test is intimately tied to the classical von Neumann-Morgenstern axioms of expected utility theory, and thus has a simple economic intuition. On the other hand, our test involves universal quantification

over a potentially infinite number of objects, while the test in Kubler, Selden and Wei (2014) can be reduced to universal quantification over a finite set.

We emphasize that what we mean by *test* is a method for falsifying the model with directly observable data. In other words, we say a model is *testable* if whenever data are inconsistent with the model, they can be demonstrated to be inconsistent. In this sense of the term test, a demonstration is distinct from an algorithm which would find this falsifying certificate. Hence, a test in our sense is not intended to be useful from a computational perspective, and as far as we can tell, ours is not in general. Indeed; there are already practical algorithms for determining when the expected utility model is falsified in our context. Rather, such a test is important for understanding the economic content of the model, by specifying a condition stated in terms of data alone, which does not reference unobservable concepts such as utilities or marginal rates of substitution. As a point of comparison, the work of Richter (1966) can be understood as providing the testable restrictions of the preference maximization hypothesis; however, no general algorithm would exist in Richter's case either.¹

Our test is perhaps most closely related to an early revealed preference test of expected utility due to Fishburn (1975). Fishburn constructs a test for an abstract environment of choice over lotteries with finite support. In his setting, one observes a finite set of binary comparisons; some are weak, and some are strict. Fishburn provides necessary and sufficient conditions for there to exist an expected utility ranking which extends the observed binary comparisons. Imagine that we observe lottery l_k weakly preferred to lottery l'_k for $k = 1, \dots, g$, and l_k strictly preferred to l'_k for $k = g + 1, \dots, K$. Fishburn establishes that these observations are consistent with expected utility maximization if there is no probability distribution over $\{1, \dots, K\}$ which puts positive probability on $\{g + 1, \dots, K\}$, and for which the mixture of the l_k 's under this probability distribution is equal to the mixture of the l'_k 's. Fishburn's test can be viewed as claiming that the smallest possible extension of the observed relations satisfying both independence and transitivity

¹In the special case where budgets are given by linear inequalities and preference satisfies monotonicity, an algorithm exists for Richter's test, namely the Afriat test. Here we refer to the abstract budget environment.

leads to no contradiction. We stress that Fishburn's test also presents with no algorithm: no recipe is given for finding the probability distribution.

In our case, we have n commodities, and a finite set of states $\Omega = \{\omega | 1, 2, \dots, S\}$. We observe a finite list of prices and contingent consumption bundles chosen at those prices (x^k, p^k) , $k \in \{1, \dots, K\}$. Consumption in state ω at observation k is of the form $x_{\omega}^k \in \mathbb{R}_+^n$. Probabilities over Ω are known and are given by the full support distribution π .

We first ask: What could reveal a violation of the joint hypothesis of expected utility and risk aversion in this context? There are only a finite set of states of the world, with known probabilities, but if the choices *were* rationalizable by an expected utility preference, there would be a natural extension to a preference over the set of all simple lotteries. One such violation would look like the following: suppose that for each x^k , there is some y^k which is feasible at prices p^k . In other words, the induced lottery l_{x^k} is *revealed preferred* to the induced lottery l_{y^k} . And suppose that there is some g for which y^g is strictly cheaper than x^g at prices p^g . In other words, the induced lottery l_{x^g} is *revealed strictly preferred* to the induced lottery l_{y^g} . Now, suppose we can find, for each k , a lottery l'_k which is a mean-preserving spread of l_{y^k} . If the data were rationalizable by a risk-averse expected utility preference, the lottery l_{x^k} would be preferred to l'_k for all k (and l_{x^g} would be strictly preferred to l'_g).

We now have a set of K pairs of lotteries (l_{x^k}, l'_k) which could be obtained in the preceding fashion. These data can be tested with Fishburn's condition. If, in fact, they violate Fishburn's condition, then we know that the original data cannot be expected utility rationalizable.

So far this is very simple. However, in the demand setting, for each observation (p^k, x^k) , there are usually infinitely many candidates for the above y^k , and for each y^k , an infinite number of possible mean-preserving spreads l'_k . This would result in an infinite number of possible $\{(l_{x^k}, l'_k)\}_{k=1}^K$ sets. While the Fishburn condition is sufficient to ensure each $\{(l_{x^k}, l'_k)\}_{k=1}^K$ set has its own preference extension, it has nothing to say about whether or not there is a *single* preference extension for the infinitely many revealed preference relations.

In fact, what we show is the following: If the data are not risk-averse expected utility

rationalizable, then there exists at least one set, $\{(l_{x^k}, l'_k)\}_{k=1}^K$, as above, that violates Fishburn's condition. In addition, they can be chosen to violate Fishburn's condition in a very stark way: one must only test the uniform lottery over $\{1, \dots, K\}$.

Moreover, the support of each l'_k can be chosen to consist only of consumption that was actually observed demanded at some state; *i.e.* the support can be chosen amongst elements of the form x_ω^k . This resonates with the idea from Polisson et al. (2015), who observe that in order to rationalize data, it is both necessary and *sufficient* to maintain consistency on the set of minimally extended “imaginary” data, constructed from those actually observed. However, while Polisson et al. (2015) is concerned with developing Afriat-style algorithms (see Afriat (1967)) for testing decision models with money lotteries, our focus is developing universal statements about data from lotteries of general consumption bundles, which provides direct falsification of the expected utility model under risk aversion.

It is important to note that due to the infinite nature of our test, our contribution lies not in providing a procedure to be implemented to check actual data; for such a test, the readers are directed to the work by Green and Srivastava (1986). Instead, the main contribution of our test is that it extends the intuition of the Fishburn test to demand-based observations: whenever the smallest possible extension of the observed relations satisfying both independence and transitivity leads to no contradiction, the data are rationalizable by risk-averse expected utility preference. In addition, the test by Green and Srivastava involves theoretical objects that are not directly observable, while our conditions directly characterize exactly which types of data are ruled out by the hypothesis of expected utility maximization, and thus can be interpreted as its UNCAF axiomatization, when observations are made in a demand-based framework.²

The idea of the proof is remarkably simple, and is a simple restatement of the dual set of linear inequalities stemming from the Afriat-style inequalities of Green and Srivastava (1986) or Varian (1983).

²UNCAF stands for *universal negation of conjunction of atomic formulas*. Chambers, Echenique and Shmaya (2014) demonstrate that theories which make no non-empirical predictions are exactly those which have UNCAF axiomatizations.

A host of other interesting papers have recently studied choice data in the context of expected utility maximization. In particular, Echenique and Saito (2015) investigates the subjective expected utility version of the model, which forms a kind of analogue of the Kubler, Selden and Wei (2014) test. It would be interesting to propose a test of our structure in the subjective expected utility framework. Epstein (2000) investigates the empirical content of the notion of probabilistic sophistication (due to Machina and Schmeidler (1992)), providing a test which can refute the hypothesis.

3.2 The Model

We assume that there is a finite state space $\Omega = \{\omega | 1, 2, \dots, S\}$ and a finite collection of consumption goods, labeled $1, 2, \dots, N$. The agent is given an objective probability distribution over states $\pi \in \Delta(\Omega)$, where for all $\omega \in \Omega$, $Pr(\omega) = \pi_\omega > 0$. An observation is a pair (p, x) , where $p \in \mathbb{R}_{++}^{SN}$ is a list of the prices of all N consumption goods under all S possible states, and $x \in \mathbb{R}_+^{SN}$ details the purchased amount of each consumption good under each state of the world.³ We assume that our *data set* \mathcal{D} consists of a K tuple of (x, p) pairs, i.e. $\mathcal{D} = \{(x^k, p^k)_{k=1}^K\}$. K is assumed finite.

In particular,

$$x^k = \begin{bmatrix} x_1^k \\ \vdots \\ x_\omega^k \\ \vdots \\ x_S^k \end{bmatrix} \quad p^k = \begin{bmatrix} p_1^k \\ \vdots \\ p_\omega^k \\ \vdots \\ p_S^k \end{bmatrix}$$

³As usual, \mathbb{R}_{++} denotes the positive reals, and \mathbb{R}_+ the nonnegative reals.

and

$$x_{\omega}^k = \begin{bmatrix} x_{\omega,1}^k \\ \vdots \\ x_{\omega,N}^k \end{bmatrix} \quad p_{\omega}^k = \begin{bmatrix} p_{\omega,1}^k \\ \vdots \\ p_{\omega,N}^k \end{bmatrix}$$

where for all ω, k, n , $x_{\omega,n}^k \geq 0$ and $p_{\omega,n}^k > 0$. Each x^k is referred to as a *contingent consumption bundle*, and x_{ω}^k a *state-specific consumption bundle*. We use $C = \mathbb{R}_+^{NS}$ to denote the set of all contingent consumption bundles.

We say that \mathcal{D} is *risk-averse expected utility rationalizable* if there exists a concave, continuous, and increasing $u : \mathbb{R}_+^N \rightarrow \mathbb{R}$ for which for all k , x^k solves

$$\max_{x \in \mathbb{R}_+^{SN}} \sum_{\omega} \pi_{\omega} u(x_{\omega})$$

subject to $p^k \cdot x \leq p^k \cdot x^k$.⁴

Given a data set \mathcal{D} , we collect all the state-specific consumption bundles x_{ω}^k observed in the data:

$$\mathcal{X} = \{x \in \mathbb{R}_+^N \mid x = x_{\omega}^k \text{ for some } k \text{ and } \omega \text{ where } (x^k, p^k) \in \mathcal{D}\}.$$

Denote the set of all simple lotteries on \mathbb{R}_+^N with finite support by $\Delta_s(\mathbb{R}_+^N)$. Denote the set of all lotteries on \mathcal{X} by $\Delta(\mathcal{X})$. Note that $\Delta(\mathcal{X}) \subseteq \Delta_s(\mathbb{R}_+^N)$.

Any contingent consumption bundle $x^k \in C$ induces an element $l_{x^k} \in \Delta(\mathcal{X})$, which places probability π_{ω} on x_{ω}^k . As such, a pair of revealed preference relations \succeq^C and \succ^C can be defined on $\Delta(\mathcal{X})$:

For $x, y \in C$, $l_x \succeq^C l_y$ if $x = x^k$ for some $(x^k, p^k) \in \mathcal{D}$ and $p^k \cdot y \leq p^k \cdot x$. For $x, y \in C$, $l_x \succ^C l_y$ if $x = x^k$ for some $(x^k, p^k) \in \mathcal{D}$ and $p^k \cdot y < p^k \cdot x$. \succeq^C is intended to represent a revealed weak preference and \succ^C a revealed strict preference.

⁴We take increasing to mean that if $x \geq y$ and $x \neq y$, then $u(x) > u(y)$.

Moreover, to test the hypothesis of risk aversion, it is natural to extend the above revealed preference relations to $\Delta_s(\mathbb{R}_+^N)$. For example, suppose that $l_x \succeq^C l_y$, and $l \in \Delta_s(\mathbb{R}_+^N)$ can be obtained by a sequence of mean-preserving spreads of l_y .⁵ If our decision maker's behavior is consistent with risk-averse expected utility maximization, it follows that l_x should also be preferred to l . These ideas motivate the following definitions.

For $l, l' \in \Delta_s(\mathbb{R}_+^N)$, $l \succeq^{m.p.s.} l'$ if l' can be obtained by a series of mean-preserving spreads of l . Define the pair of binary relations \succeq^R and \succ^R on $\Delta_s(\mathbb{R}_+^N)$ by

$$l \succeq^R l'' \text{ if there exists } l' \text{ such that } l \succeq^C l' \succeq^{m.p.s.} l''$$

and

$$l \succ^R l'' \text{ if there exists } l' \text{ such that } l \succ^C l' \succeq^{m.p.s.} l''$$

If the agent's behavior is consistent with risk-averse expected utility maximization, the pair of relations \succeq^R, \succ^R will necessarily satisfy Fishburn's condition on $\Delta_s(\mathbb{R}_+^N)$; *i.e.* if $l_k \succeq^R l'_k$ for $k = 1, \dots, g$, and $l_k \succ^R l'_k$ for $k = g + 1, \dots, K$, then there are no $\{\mu_i\}_{i=1}^K \subseteq \mathbb{R}_+^K$, with $\sum_{k=g+1}^K \mu_k > 0$, and $\sum_1^K \mu_k l_k = \sum_1^K \mu_k l'_k$. As we show in our main result, it turns out that a *sufficient* condition for the data \mathcal{D} to conform with risk aversion and expected utility maximization is that the restriction of \succeq^R, \succ^R to $\Delta(\mathcal{X})$ satisfies Fishburn's condition.

For every data set $\mathcal{D} = \{(x^k, p^k)_{k=1}^K\}$, the following are equivalent:

I For any $\{l'_k\}_{k=1}^K \subseteq \Delta(\mathcal{X})$ for which $l_{x^k} \succeq^R l'_k$ for all k , there is no $\{\mu_k\}_{k=1}^K \subseteq \mathbb{R}_+^K$ for which $\sum_{\{k: l_k \succ^R l'_k\}} \mu_k > 0$ and $\sum_1^K \mu_k l_{x^k} = \sum_1^K \mu_k l'_k$.

II Suppose that for each $k \in \{1, \dots, K\}$ and $\omega \in \Omega$, $S_\omega^k : \{1, \dots, K\} \times \Omega \rightarrow \mathbb{R}_+$ is a function,

⁵That is, if there exists a random variable ε such that $l \stackrel{d}{=} l_y + \varepsilon$ with $E(\varepsilon|l_y) = 0$. “ $\stackrel{d}{=}$ ” here means “has the same distribution as”. See Rothschild and Stiglitz (1970) for more details.

such that for all $k, \omega, \sum_{g,\tau} S_{\omega}^k(g, \tau) = \pi_{\omega} = \sum_{g,\tau} S_{\tau}^g(k, \omega)$. If, in addition, for all k ,

$$p^k \cdot x^k \geq p^k \cdot \left(\frac{\sum_g \sum_{\tau} S_{\omega}^k(g, \tau) x_{\tau}^g}{\pi_{\omega}} \right)_{\omega \in \Omega}$$

then there is no k for which $p^k \cdot x^k > p^k \cdot \left(\frac{\sum_g \sum_{\tau} S_{\omega}^k(g, \tau) x_{\tau}^g}{\pi_{\omega}} \right)_{\omega \in \Omega}$.⁶

III For all $\omega, \tau \in \Omega$ and $k, g \in \{1, \dots, K\}$ there exist $u_{\omega}^k, u_{\tau}^g \geq 0$ and $\lambda_k, \lambda_g > 0$ s.t. $u_{\omega}^k \leq u_{\tau}^g + \lambda_g \frac{p_{\tau}^g}{\pi_{\tau}} \cdot (x_{\omega}^k - x_{\tau}^g)$.⁷

IV Data set \mathcal{D} is risk-averse expected utility rationalizable.

Before proceeding, we comment on cases I and II, which are our contribution. Case I considers the smallest possible preference extension “consistent” with the data, risk-aversion, and the expected utility hypothesis. It claims that if this extension is meaningfully defined; in that we cannot derive that a lottery l is strictly preferred to itself, then the data are expected utility rationalizable. Importantly, we only need to consider lotteries whose support are actual observed consumption bundles. This can be seen as a natural analogue of Fishburn’s condition as applied to l_{x^k} and l'_k .

Case II demonstrates a dual system of linear inequalities to the inequalities of case III, which was derived previously by Green and Srivastava (1986). The interpretation of the terms S_{ω}^k is as a system of probability weights. To obtain some intuition on Case II, suppose that the inequalities therein are satisfied, then one can find a contradiction as follows: For each k , by demand behavior, the inequalities in Case II imply that the lottery l_{y^k} induced by the contingent consumption bundle $\left(\frac{\sum_g \sum_{\tau} S_{\omega}^k(g, \tau) x_{\tau}^g}{\pi_{\omega}} \right)_{\omega \in \Omega}$ is revealed weakly worse than the lottery l_{x^k} induced

$$6 \left(\frac{\sum_g \sum_{\tau} S_{\omega}^k(g, \tau) x_{\tau}^g}{\pi_{\omega}} \right)_{\omega \in \Omega} = \left(\frac{\sum_g \sum_{\tau} S_1^k(g, \tau) x_{\tau}^g}{\pi_1}, \dots, \frac{\sum_g \sum_{\tau} S_S^k(g, \tau) x_{\tau}^g}{\pi_S} \right) \text{ i.e. } \frac{\sum_g \sum_{\tau} S_{\omega}^k(g, \tau) x_{\tau}^g}{\pi_{\omega}} \text{ is the consumption in state } \omega.$$

⁶Green and Srivastava’s proof of this statement assumes the non-emptiness of u ’s superdifferential over \mathbb{R}_+^n ; however, it is easy to modify their proof even with empty superdifferential on the boundary. Essentially, whenever x^g is known to be a utility maximizer, we can always find $\nabla u(x_{\tau}^g)$ in the superdifferential of u for which $\nabla u(x_{\tau}^g) = \lambda_g \frac{p_{\tau}^g}{\pi_{\tau}}$ (see Theorem 28.3 in Rockafellar (1997)). So $u_{\omega}^k \leq u_{\tau}^g + \nabla u(x_{\tau}^g) \cdot (x_{\omega}^k - x_{\tau}^g) = u_{\tau}^g + \lambda_g \frac{p_{\tau}^g}{\pi_{\tau}} \cdot (x_{\omega}^k - x_{\tau}^g)$.

by x^k , with strict preference for at least one k . Observe that l_{y^k} is a lottery that places probability π_ω on $\frac{\sum_g \sum_\tau S_\omega^k(g, \tau) x_\tau^g}{\pi_\omega}$. Since $\sum_{g, \tau} S_\omega^k(g, \tau) = \pi_\omega$, simple algebra (included in the proof) shows that the lottery l'_k , which places probability weight $S_\omega^k(g, \tau)$ on x_τ^g , is a mean-preserving spread of l_{y^k} . If the data were really consistent with the hypothesis of *risk-averse* expected utility maximization, transitivity would imply that for each k , the lottery l'_k should be worse than the lottery l_{x^k} , strictly so for at least one k . We now have in total K revealed preference relations between the pairs of lotteries l_{x^k} and l'_k . As we demonstrate in the proof, the condition $\sum_{g, \tau} S_\tau^g(k, \omega) = \pi_\omega$ then allows us to find a violation by applying the condition from Fishburn (1975) on the lotteries l_{x^k} and l'_k across all k .

The following example illustrates the theorem.

Example 1 Consider the case $k \in \{1, 2\}$, $\Omega = \{1, 2\}$ and $N = 2$: There are 2 observations, each consisting of the price and purchased quantity for the consumptions good under 2 possible states of the world. Suppose each of the two states are equally likely; $\pi_1 = \pi_2 = .5$. Suppose we observe:

$$(x^1, p^1) = \left(\begin{array}{c} \begin{bmatrix} 0 \\ 0 \\ 10 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 10 \\ 5 \\ 10 \end{bmatrix} \end{array} \right) \quad (x^2, p^2) = \left(\begin{array}{c} \begin{bmatrix} 4 \\ 2 \\ 6 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \\ 5 \\ 10 \end{bmatrix} \end{array} \right)$$

In this case there is no violation of GARP. However, since the hypothesis of risk-averse EU preference is stronger than than GARP, we show that this case still violates our conditions.

Violation of Statement I: The induced lotteries by x^1 and x^2 are $l_{x^1} = ((10, 5), 1/2; (0, 0), 1/2)$ and $l_{x^2} = ((4, 2), 1/2; (6, 3), 1/2)$, respectively. To see that this is a violation of statement I, consider contingent consumption bundles $y^1 = y^2 = ((5, 2.5); (5, 2.5))$

which induce $l_{y^1} = l_{y^2} = ((5, 2.5), 1)$. Clearly $p^1 \cdot x^1 \geq p^1 \cdot y^1$, and $p^2 \cdot x^2 > p^2 \cdot y^2$. So by definition $l_{x^1} \succeq^C l_{y^1}$ and $l_{x^2} \succ^C l_{y^2}$.

Observe that the lottery $l'_1 = ((4, 2), 1/2; (6, 3), 1/2)$ is a mean-preserving spread of l_{y^1} and the lottery $l'_2 = ((10, 5), 1/2; (0, 0), 1/2)$ is a mean-preserving spread of l_{y^2} . By definition $l_{x^1} \succeq^R l'_1$ and $l_{x^2} \succ^R l'_2$. However,

$$\frac{1}{2}l_{x^1} + \frac{1}{2}l_{x^2} = \frac{1}{2}l'_1 + \frac{1}{2}l'_2$$

This constitutes a violation of Statement I.

Violation of Statement II:

Set $S_1^1(2, 1) = S_1^2(1, 1) = \frac{1}{5}$, $S_1^1(2, 2) = S_2^2(1, 1) = \frac{3}{10}$, $S_2^1(2, 1) = S_1^1(1, 2) = \frac{3}{7}$, and $S_2^1(2, 2) = S_2^2(1, 2) = \frac{1}{14}$.

To solve:

$$\begin{bmatrix} 0 \\ 0 \\ 10 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 10 \\ 5 \\ 10 \end{bmatrix} > 2 \cdot \begin{bmatrix} 5 \\ 10 \\ 5 \\ 10 \end{bmatrix} \cdot \begin{bmatrix} S_1^1(2, 1) * 4 + S_1^1(2, 2) * 6 \\ S_1^1(2, 1) * 2 + S_1^1(2, 2) * 3 \\ S_2^1(2, 1) * 4 + S_2^1(2, 2) * 6 \\ S_2^1(2, 1) * 2 + S_2^1(2, 2) * 3 \end{bmatrix}$$

$$\begin{bmatrix} 4 \\ 2 \\ 6 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 8 \\ 5 \\ 10 \end{bmatrix} \geq 2 \cdot \begin{bmatrix} 4 \\ 8 \\ 5 \\ 10 \end{bmatrix} \cdot \begin{bmatrix} S_1^2(1, 1) * 0 + S_1^2(1, 2) * 10 \\ S_1^2(1, 1) * 0 + S_1^2(1, 2) * 5 \\ S_2^2(1, 1) * 0 + S_2^2(1, 2) * 10 \\ S_2^2(1, 1) * 0 + S_2^2(1, 2) * 5 \end{bmatrix}$$

$$S_1^1(2, 1) + S_1^1(2, 2) = S_2^1(2, 1) + S_2^1(2, 2) = \frac{1}{2}$$

A couple of observations are in order. It can be shown that both (I) and (II) of our properties imply GARP. Suppose by means of contradiction that GARP is violated, *i.e.* that there are contingent consumption bundles z^{k_1}, \dots, z^{k_m} such that $p^{k_1} \cdot z^{k_1} \geq p^{k_1} \cdot z^{k_2}, p^{k_2} \cdot z^{k_2} \geq p^{k_2} \cdot z^{k_3}, \dots, p^{k_m} \cdot z^{k_m} > p^{k_m} \cdot z^{k_1}$, where without loss we may assume there is no repetition in the cycle. This implies $l_{z^{k_1}} \succ^C l_{z^{k_2}} \succ^C \dots \succ^C l_{z^{k_m}} \succ^C l_{z^{k_1}}$.

To see that (I) implies GARP, observe that since \succ^C implies \succ^R and \succ^C implies \succ^R , we have $l_{z^{k_1}} \succ^R l_{z^{k_2}} \succ^R \dots \succ^R l_{z^{k_m}} \succ^R l_{z^{k_1}}$. Let $l_{x^i} = l_{z^{k_i}}$ and $l'_i = l_{z^{k_{i+1}}}$ as in property (I), then a uniform distribution μ over the indices $i = 1, 2, \dots, m$ constitutes a violation of (I).

For (II), consider the following set of $S_\omega^k(g, \tau)$'s in property II: For $k = k_i$ for some i (that is, if k shows up in the cycle)

$$S_\omega^{k_i}(g, \tau) = \begin{cases} \pi_\omega & \text{if } g = k_{i+1} \text{ and } \tau = \omega \\ 0 & \text{otherwise} \end{cases}$$

and for $k \neq k_i$ for any i (k not in the cycle)

$$S_\omega^k(g, \tau) = \begin{cases} \pi_\omega & \text{if } g = k \text{ and } \tau = \omega \\ 0 & \text{otherwise} \end{cases}$$

Then the cycle condition gives a violation of property (II), a contradiction.

Finally, we wish to emphasize that the result is by no means a trivial consequence of Fishburn (1975). In his paper, he also considers the issue of testing the consistency of revealed preference relations with functional restrictions on the von Neumann-Morgenstern utility index (as we wish to test for concavity and monotonicity). Specifically, he wants to test when observed data are consistent with the utility index u belonging to some convex cone

\mathcal{U} . Again, he assumes a finite number of relations (which does not hold in our context). A natural guess is that if l_k is revealed weakly preferred to l'_k for $k = 1, \dots, g$ and revealed strictly preferred to l'_k for $k = g + 1, \dots, K$, then if there is $\mu \in \Delta(K)$ for which $\mu(\{g + 1, \dots, K\}) > 0$ and $u \cdot (\sum_k \mu_k l'_k) \geq u \cdot (\sum_k \mu_k l_k)$ for all $u \in \mathcal{U}$, then the observed data are inconsistent with expected utility maximization with utility index $u \in \mathcal{U}$ ⁸. In our case, for example, we would consider the cone of concave, nondecreasing and locally non-satiated functions; the claim would then be that $\sum_k \mu_k l'_k$ second order stochastically dominates $\sum_k \mu_k l_k$. Of course, the existence of such a μ refutes the hypothesis of expected utility rationalization with $u \in \mathcal{U}$, but for technical reasons, the converse statement need not hold in general (it would hold, for example, if the cone \mathcal{U} were polyhedral, which is not the case here). However, we are able to show that owing to the special structure of linear pricing, a converse statement along the lines of this idea does in fact hold in the demand-based environment. In fact, it holds *even though observed revealed preference relations are infinite*.

Proof.

(III \Leftrightarrow IV)

The equivalence of III and IV is due to Green and Srivastava (1986).

⁸Here we continue to use x and z for lotteries, and dot product for integration with respect to measures.

(II \Leftrightarrow III)

We proceed to show that II and III are equivalent. To this end, observe that III does not hold if and only if there is no solution to the following linear system.⁹ $Ab \geq 0$ and $\lambda \gg 0$, where

$$b = \begin{bmatrix} u_1^1 \\ u_2^1 \\ \vdots \\ u_S^K \\ \lambda \end{bmatrix} \quad \lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_K \end{bmatrix}$$

and A is equal to the top two quadrants of the matrix below:

$$T = \begin{array}{c} \eta_{1,1,1,1} \\ \vdots \\ \eta_{k,\omega,g,\tau} \\ \vdots \\ \eta'_k \\ \vdots \end{array} \begin{array}{c} u_1^1 \quad \dots \quad u_\omega^k \quad \dots \quad u_\tau^g \quad \dots \quad u_S^K \\ \left[\begin{array}{cccccc|cccc} 0 & \dots & 0 & \dots & 0 & \dots & 0 & \lambda_1 & \dots & \lambda_k & \dots & \lambda_K \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & -1 & \dots & 0 & 0 & \dots & \frac{p_\omega^k}{\pi_k} \cdot (x_\tau^g - x_\omega^k) & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \end{array} \right] \end{array}$$

By construction of T and a standard theorem of the alternative (see for example Mangasarian (1994) p. 30), the nonexistence of b, λ such that $Ab \geq 0$ and $\lambda \gg 0$, is equivalent to the existence of $\eta \geq 0$ such that $T'\eta \leq 0$, where

⁹Vector inequalities are $x \geq y$ if $x_i \geq y_i$ for all i and $x \gg y$ if $x_i > y_i$ for all i .

$$\eta = \begin{bmatrix} \eta_{1,1,1,1} \\ \vdots \\ \eta_{K,S,K,S} \\ \eta' \end{bmatrix} \quad \eta' = \begin{bmatrix} \eta'_1 \\ \vdots \\ \eta'_K \end{bmatrix}$$

such that at least one $\eta'_k > 0$.

This is equivalent to

$$\sum_{\omega} \sum_{(g,\tau) \neq (k,\omega)} \eta_{k,\omega,g,\tau} \frac{p_{\omega}^k}{\pi_{\omega}} \cdot (x_{\tau}^g - x_{\omega}^k) \leq 0 \quad \forall k \quad (3.1)$$

with strict inequality for at least one k , and

$$\sum_{(g,\tau) \neq (k,\omega)} \eta_{k,\omega,g,\tau} = \sum_{(g,\tau) \neq (k,\omega)} \eta_{g,\tau,k,\omega} \quad \forall k, \omega \quad (3.2)$$

We claim that a solution to systems (3.1) and (3.2), implies the existence of $\gamma_{k,\omega,g,\tau} \geq 0$ so that

$$\sum_{\omega} \sum_{(g,\tau)} \gamma_{k,\omega,g,\tau} \frac{p_{\omega}^k}{\pi_{\omega}} \cdot (x_{\tau}^g - x_{\omega}^k) \leq 0 \quad \forall k \quad (3.3)$$

$$\sum_{(g,\tau)} \gamma_{k,\omega,g,\tau} = \sum_{(g,\tau)} \gamma_{g,\tau,k,\omega} = \pi_{\omega} \quad \forall k, \omega \quad (3.4)$$

with at least one inequality in (3.3) being strict, effectively showing (3.3) and (3.4) are equivalent to (3.1) and (3.2).

To see this, list the $\eta_{k,\omega,g,\tau}$'s from systems (3.1) and (3.2) as in Figure 3.1 (Notice that system (3.2) ensures that columns and rows passing through the same diagonal element, like the column and row in red and blue boxes, sum up to the same number.) We now construct a

$$\begin{array}{c}
\begin{array}{cccccc}
& & 1,1 & & 1,\omega & & 1,S & & k,\omega & & K,S \\
1,1 & \left[\begin{array}{cccccc}
\eta_{1,1,1,1} & \cdots & \eta_{1,1,1,\omega} & \cdots & \eta_{1,1,1,S} & \cdots & \eta_{1,1,k,\omega} & \cdots & \eta_{1,1,K,S} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \cdots \\
1,\omega & \eta_{1,\omega,1,1} & \cdots & \eta_{1,\omega,1,\omega} & \cdots & \eta_{1,\omega,1,S} & \cdots & \eta_{1,\omega,k,\omega} & \cdots & \eta_{1,\omega,K,S} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \cdots \\
1,S & \eta_{1,S,1,1} & \cdots & \eta_{1,S,1,\omega} & \cdots & \eta_{1,S,1,S} & \cdots & \eta_{1,S,k,\omega} & \cdots & \eta_{1,S,K,S} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \cdots \\
k,\omega & \eta_{k,\omega,1,1} & \cdots & \eta_{k,\omega,1,\omega} & \cdots & \eta_{k,\omega,1,S} & \cdots & \eta_{k,\omega,k,\omega} & \cdots & \eta_{k,\omega,K,S} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \cdots \\
K,S & \eta_{K,S,1,1} & \cdots & \eta_{K,S,1,\omega} & \cdots & \eta_{K,S,1,S} & \cdots & \eta_{K,S,k,\omega} & \cdots & \eta_{K,S,K,S}
\end{array} \right.
\end{array}
\end{array}$$

Figure 3.1: η matrix

new matrix, say, Λ , with generic element $\lambda_{k,\omega,g,\tau}$ by raising all diagonal entries of the η matrix, leaving all remaining entries the same, so that there is some $M > 0$ for which $\sum_{(g,\tau)} \eta_{k,\omega,g,\tau} = \sum_{(g,\tau)} \eta_{g,\tau,k,\omega} = M\pi_\omega$.¹⁰ Since $p_\omega^k \cdot (x_\omega^k - x_\omega^k) = 0$, and since the diagonal element shows up both in the column and row, the resulting η matrix satisfies (3.3) (with λ 's in place of η 's), and the first equality in system (3.4). Finally, the γ terms are constructed by dividing each element of the matrix Λ by M .

Rearranging inequalities (3.3) gives

$$\begin{aligned}
\sum_{\omega} \sum_g \sum_{\tau} \gamma_{k,\omega,g,\tau} \frac{p_\omega^k}{\pi_\omega} \cdot (x_\tau^g - x_\omega^k) &= \sum_{\omega} p_\omega^k \cdot \left(\sum_g \sum_{\tau} \frac{\gamma_{k,\omega,g,\tau} x_\tau^g}{\pi_\omega} - \sum_g \sum_{\tau} \frac{\gamma_{k,\omega,g,\tau} x_\omega^k}{\pi_\omega} \right) \\
&= \sum_{\omega} p_\omega^k \cdot \left(\sum_g \sum_{\tau} \frac{\gamma_{k,\omega,g,\tau} x_\tau^g}{\pi_\omega} - x_\omega^k \right) \leq 0
\end{aligned}$$

with at least one strict inequality. The second equality follows from (3.4). This together with (3.4) establishes the equivalence of II and III, by taking $S_\omega^k(g, \tau) = \gamma_{k,\omega,g,\tau}$.

¹⁰One simple way of doing this is to pick M large enough so that $\min_{\omega} \pi_{\omega} M > \max_{\omega} \sum_{(g,\tau) \neq (k,\omega)} \eta_{k,\omega,g,\tau}$.

(IV \Rightarrow I)

That IV implies I is straightforward. Let $u : \mathbb{R}_+^N \rightarrow \mathbb{R}$ be any concave, nondecreasing and locally non-satiated utility function. For lottery l , let $u \cdot l$ denote the expected utility of l , $\sum_{x \in l} l(x)u(x)$.

Suppose that \mathcal{D} is risk-averse expected utility rationalizable by u , and suppose by means of contradiction that statement I is not true

For all $l'_k \in \Delta_g(\mathbb{R}_+^N)$, $l_{x^k} \succeq^R l'_k$ implies $u \cdot l_{x^k} \geq u \cdot l'_k$, and $l_{x^k} \succ^R l'_k$ implies $u \cdot l_{x^k} > u \cdot l'_k$. Since expected utility is linear in lottery mixtures, we have that $u \cdot (\sum_1^K \mu_k l_{x^k}) > u \cdot (\sum_1^K \mu_k l'_k)$, a contradiction to $\sum_1^K \mu_k l_{x^k} = \sum_1^K \mu_k l'_k$.

(I \Rightarrow II)

We now show that I implies II. Suppose by means of contradiction that there is a solution to the system listed in II. We will show that this implies I is false. Let

$$y^k = \left(\frac{\sum_g \sum_\tau S_\omega^k(g, \tau) x_\tau^g}{\pi_\omega} \right)_{\omega \in \Omega}$$

By II, we have $p^k \cdot x^k \geq p^k \cdot y^k \forall k$ with $>$ for at least one k . By definition of \succeq^C , $l_{x^k} \succeq^C l_{y^k}$, with \succ^C for at least one k .

Next, observe that l_{y^k} places probability π_ω at $\frac{\sum_g \sum_\tau S_\omega^k(g, \tau) x_\tau^g}{\pi_\omega}$ for each ω . Let l'_k be the lottery that puts probability $\sum_\omega S_\omega^k(g, \tau)$ on x_τ^g . Since $\sum_{g, \tau} S_\omega^k(g, \tau) = \pi_\omega$, l'_k can be obtained from l_{y^k} by spreading, for each ω , the probability π_ω placed on $\frac{\sum_g \sum_\tau S_\omega^k(g, \tau) x_\tau^g}{\pi_\omega}$ to probabilities $S_\omega^k(g, \tau)$'s on x_τ^g 's, $(g, \tau) \in \{1, \dots, K\} \times \Omega$. Moreover, $\frac{\sum_g \sum_\tau S_\omega^k(g, \tau) x_\tau^g}{\pi_\omega}$ is a weighted average of the x_τ^g 's by weights $S_\omega^k(g, \tau)$'s. So for each ω the spread described above is a mean-preserving spread in the sense of Rothschild and Stiglitz (1970), and l'_k can be obtained from l_{y^k} by a finite number of mean-preserving spread.

By definition of \succeq^R , we have obtained lotteries l_{x^k} and l'_k such that $l_{x^k} \succeq^R l'_k \forall k$, with \succ^R

for at least one k . In order to contradict I, it only remains now to find $\{\mu_k\}_{k=1}^K$ such that $\mu_k \geq 0$, $\sum_{\{k:l_k > Rl'_k\}} \mu_k > 0$ and $\sum_1^K \mu_k l_{x^k} = \sum_1^K \mu_k l'_k$. As it turns out, it suffices to take $\mu_k = \frac{1}{K}$ for each k :

The lottery $\sum_{k=1}^K \frac{1}{K} l'_k$ places probability $\frac{1}{K} \sum_k \sum_{\omega} S_{\omega}^k(g, \tau) = \frac{\pi_{\tau}}{K}$ on each x_{τ}^g , $(g, \tau) \in \{1, \dots, K\} \times \Omega$, while the lottery $\sum_{k=1}^K \frac{1}{K} l_{x^k}$, places $\frac{\pi_{\tau}}{K}$ on each x_{τ}^g . So $\sum_{k=1}^K \frac{1}{K} l'_k = \sum_{k=1}^K \frac{1}{K} l_{x^k}$. This constitutes a contradiction to I (in particular, the contradiction comes in the form of a uniform distribution over the observations $1, \dots, K$).

■

3.3 Conclusion

We have developed a universal test for the risk-averse expected utility environment with many commodities. Of interest for future research would be an analogous test in the subjective expected utility context, following the work of Echenique and Saito (2015). The difficulty inherent in this approach rests in the fact that the inequalities in III of Theorem 3.2 are polynomial, rather than linear. While we have some conjectures on what might be an appropriate test, these are very speculative.

A final remark is in order. Observe that when $|\Omega| = 1$ (and hence $\pi_{\omega} = 1$ for ω for which $\Omega = \{\omega\}$), we are back to the environment of Afriat (1967). In such an environment, the function S referenced in Theorem 3.2, condition II can be taken to be a function of $\{1, \dots, K\}$ alone. And condition II in this case tells us that $\sum_k S_k(l) = \sum_k S_l(k) = 1$ for each l ; in other words, viewing S as a matrix, the matrix is *bistochastic*. Now, one of the contributions of Afriat (1967) is that condition II is necessary and sufficient for concave rationalization when the matrix S is restricted to be a *permutation* matrix; that is, a matrix consisting solely of zeroes and ones. Of course, it is well-known that the permutation matrices are the extreme points of the set of bistochastic matrices (this is the celebrated theorem of Birkhoff (1946) and von Neumann (1953)). A natural conjecture is that a similar statement may hold here; that it is enough to check the extreme points of the set of S functions satisfying condition II of Theorem 3.2.

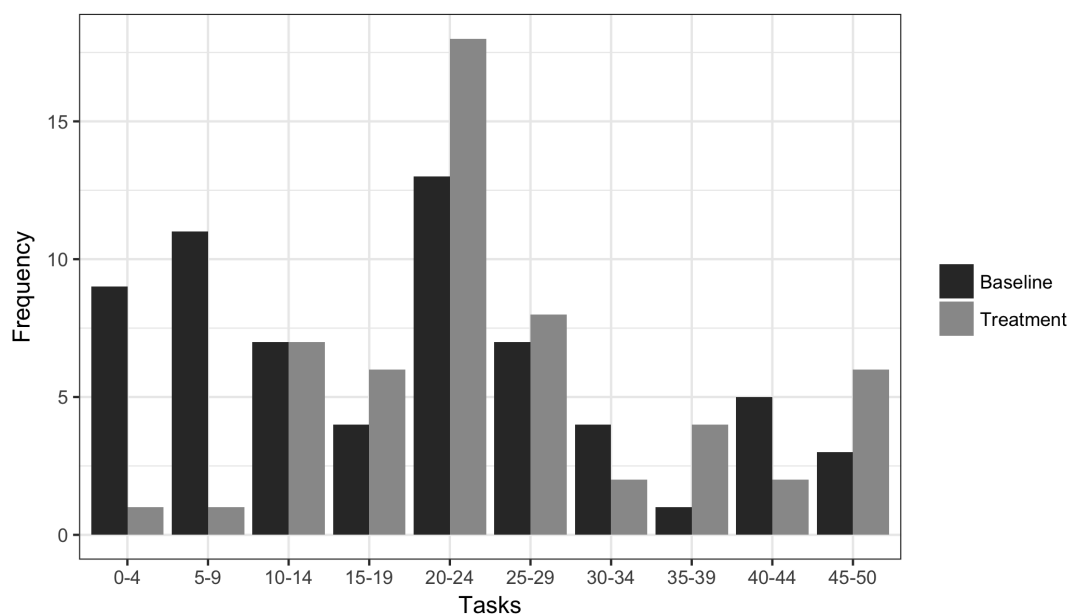
Acknowledgements

Chapter 3, in full, has been published. Martinez, Seung-Keun; Chambers, Christopher; Liu, Ce. “A Test for Risk-Averse Expected Utility.” The dissertation author was one of three primary investigators and authors of this material.

Appendix A

Social Comparisons in Peer Effects

A.1 Results



Notes: This figure shows that the difference in task choice between treatment and baseline arises from subjects who were willing to do 20 or fewer tasks. Given the established correlation between choice and beliefs, this figure suggests that the refined signal corrected beliefs so that more subjects believed the high signal and, consequently, were willing to do more tasks.

Figure A.1: Distribution of Task Choice by Treatment

Table A.1: Distribution of Task Choice by Treatment

DV: Tasks Choice	
\$5	3.234*** (0.472)
\$10	7.468*** (0.741)
\$15	11.270*** (0.957)
\$20	14.541*** (1.185)
Treat x \$2	0.559 (1.098)
Treat x \$5	1.538 (1.293)
Treat x \$10	2.433 (1.572)
Treat x \$15	3.020* (1.779)
Treat x \$20	4.537** (2.038)
Constant	6.450*** (0.815)
R-Squared	0.212
Subjects	219
Observations	1095

Notes: This table individually presents the results of Figure 1.5 for each donation amount. Robust standard errors, clustered at the subject level, presented in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Task Choice by Belief, Signal, Treatment

	DV: Task Choice	
	(1)	(2)
High Belief	12.736*** (2.151)	11.931*** (2.599)
High Signal		1.420 (2.504)
Treatment		1.018 (2.642)
Constant	12.524*** (1.638)	11.596*** (1.673)
R-Squared	0.216	0.221
Observations	119	119

Notes: We regress task choice on an indicator that the subject believes the high signal is true. Robust standard errors presented in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Effect of Treatment on Likelihood of Choosing At Least Twenty Tasks

	DV: $1\{\text{Tasks} \geq 20\}$
Treatment	0.212** (0.087)
Constant	0.516*** (0.063)
R-Squared	0.047
Observations	119

Notes: We regress an indicator for choosing to do 20 or more tasks on treatment. Robust standard errors in parantheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Quantile Regressions

	DV: Task Choice				
	(1) 20th	(2) 40th	(3) 50th	(4) 60th	(5) 80th
Treatment	10.000*** (2.434)	10.000*** (3.099)	0.000 (3.045)	3.000 (3.068)	5.000 (6.612)
Constant	5.000*** (1.864)	10.000*** (2.438)	20.000*** (2.365)	20.000*** (2.502)	30.000*** (3.523)
R-Squared	0.050	0.050	.	0.050	0.050
Observations	119	119	119	119	119

Notes: This table quantifies the distributional difference in task choice shown in Figure A.1. Standard errors in parantheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Quantile Regressions

	DV: Task Choice				
	(1) 10th	(2) 30th	(3) 50th	(4) 70th	(5) 90th
Treatment	8.000** (3.761)	12.000*** (2.967)	0.000 (3.045)	0.000 (3.376)	5.000 (6.777)
Constant	2.000 (1.981)	8.000*** (2.046)	20.000*** (2.365)	25.000*** (2.621)	40.000*** (5.485)
R-Squared	0.050	0.050	.	.	0.050
Observations	119	119	119	119	119

Notes: Standard errors in parantheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Small Sample t-test:
Effect of High Signal on Task
Choice

	Task Difference
High Signal	9.608** (3.357)
N High Signal	51
N Low Signal	4
N Overall	55

Notes: Standard errors in parantheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Effect of Treatment on Belief
of Received Signal

	DV: 1{Believe Signal}
Treatment	0.341*** (0.078)
Constant	0.531*** (0.063)
R-Squared	0.135
Observations	119

Notes: Standard errors in parantheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.2 Experiment Instructions

All experiment sessions were run using Otree. The instructions were shown on a step-by-step basis with each feature of the experiment explained on a sequence of screens.

Instructions: Anticipation Experiment

Screen 1

Hello and welcome, today you will be participating in an experiment on economic decision making. Funds for this experiment have been provided by the University of California. You will be paid for your participation in this experiment. Your final payment will consist of a fixed \$5 show-up payment and a \$10 completion bonus. You will be paid privately in cash after the experiment has concluded. We ask that you please silence and put away all cell phones and any other personal electronic devices now and for the remainder of the experiment.

In this experiment you will be presented with an opportunity to complete a task for charity. For this project, we are partnering with Dr. James Rolfe – founder of the Afghan Dental Relief Project – to bring modern dental care to the very poor in Afghanistan. The A.D.R.P. is a charitable organization that provides free dental services and dental health education to the poorest families and individuals in the city of Kabul, the capital of Afghanistan. 100% of the donations that you generate from this experiment will be used to purchase dental supplies and ship them to A.D.R.P.'s free dental clinic. Please take the time to carefully read the news article below that details the work Dr. Rolfe and the A.D.R.P. have accomplished and can continue to accomplish with your help.

Screen 2

Today, you will have the opportunity to complete tasks to generate donations to A.D.R.P.'s free dental clinic. The task will be to transcribe captchas. For each task, you will be shown

an image of text. Your objective is to correctly type the text that is shown to you in the space provided. Each correctly transcribed captcha will be counted as one completed task. You will have 3 chances to correctly transcribe each captcha. If you fail to correctly transcribe a captcha within three tries you will simply be shown a different captcha. You will now be shown 5 captchas to transcribe so that you are familiar with the task. Please note that each captcha may consist of both capital and lower case letters, the numbers 2-9, and special characters !, %, &, ?.

Screen 3

5 Sample Captchas

Screen 4: Tasks for Donations

For this experiment you will decide how many captchas you are willing to complete in order to generate donations for the A.D.R.P.'s free dental clinic. That is, you will choose the maximum number of captchas you are willing to do for donations of \$2, \$5, \$10, \$15, and \$20. You will be able to select between 0 and 50 tasks for each possible donation amount. For example, if for a donation of \$5 dollars you select 12 tasks, then you are indicating that you are willing to do a maximum of 12 tasks in exchange for a \$5 donation to the A.D.R.P.

After you have made your decisions we will randomly assign each person a donation amount and a number of tasks. If the the number of tasks you are assigned is fewer than the maximum number you selected for yourself, then you will complete your assigned number of tasks and we will give the assigned donation to the A.D.R.P. If the randomly assigned number of tasks is higher than your chosen number, then you will complete no tasks and no donation will be given to the A.D.R.P.

For example, suppose that for a donation of \$15 you choose to do at most 30 tasks. Then if you are randomly assigned \$15 and 20 tasks, then you will complete 20 tasks and \$15 will be given to the Afghan Dental Relief Project. On the other hand if you are randomly assigned \$15

and 40 tasks, then you will do 0 tasks and no donation will be given to the A.D.R.P.

Each person will be assigned their own number of tasks and their own donation amount. No one will learn your assigned donation amount or number of tasks. Please note that the randomization device is “fair” – all possible donation and task amounts will have equal probability of being assigned. So it is in your best interest to treat each decision as if it is the one that counts.

You will be free to leave the experiment as soon as you complete your tasks. You will be free to go immediately if you have no tasks to complete. Your completion bonus will be \$10 regardless of the number of tasks you have to do.

Comprehension Questions

Suppose you choose to do a maximum of 30 tasks for a \$10 donation. Under which of the following conditions would you complete tasks for a \$10 donation?

- If I am randomly assigned a \$10 donation and 20 tasks.
- If I am randomly assigned a \$10 donation and 40 tasks.

Suppose you are assigned a \$10 donation and 21 tasks. How many tasks will you be asked to do?

- 21 if I choose more than 21 tasks for a \$10 donation.
- 21 if I choose fewer than 21 tasks for a \$10 donation.

Suppose you are assigned a \$5 donation and 15 tasks. How many tasks will you be assigned to do?

- 0 if I choose more than 15 tasks for a \$10 donation.
- 0 if I choose fewer than 15 tasks for a \$10 donation.

Screen 5: Guessing other’s choices

In addition to selecting how many tasks you are willing to do for each possible donation amount, you will guess how your decisions compare to the decisions of previous participants.

That is, you will guess what percent of previous participants were willing to do a larger number of tasks than you for each possible donation. Specifically, you will say whether you believe that at least 95%, 75%, 50%, 25%, or 0% of all previous participants were willing to do more tasks than you.

As an example, suppose that for a donation of \$5 dollars you choose to do a maximum of 12 tasks and think that at least 50%, but fewer than 75%, of previous participants were willing to do more than 12 tasks for a \$5 donation. Then you will indicate that you believe that at least 50% of all previous experiment participants were willing to do more tasks than you for a \$5 donation.

Comprehension Question

Suppose that you choose to do at most 10 tasks for a donation of \$15 dollars and that you believe that at least 75%, but fewer than 95%, of previous participants were willing to do more than 10 tasks for a \$15 donation. How would you answer the question “What percent of previous participants were willing to do more tasks that you?”

- 95%
- 75%
- 50%
- 25%
- 0%

Screen 5B: Seeing other’s choices (Treatment Only)

Lastly, you will learn how your decisions actually compare to the decisions of all those who participated in this experiment before you. After everyone has made their decisions and guesses, you will open the yellow envelope on your desk. The contents of this envelope will show you how many tasks at least 95%, 75%, 50%, 25% and 0% of all previous participants willing to do. Therefore, you will learn the true proportion of previous participants that were

willing to do more tasks than you. You will then verify that you understand this information by answering a series of multiple choice questions. Please do not open these envelopes until you are instructed to do so.

For example, suppose for a donation of \$5 dollars you were willing to do 12 tasks. And suppose that the information in the yellow envelope shows you that 50% of previous participants were willing to do at least 15 tasks for a \$5 donation and 75% were willing to do at least 10 tasks for \$5 donation. This would mean that at least 50%, but less than 75%, of previous participants were willing to do more tasks than you for a \$5 donation.

Comprehension Question

Suppose that you choose to do at most 35 tasks for a donation of \$15 dollars. Also suppose you learn from the information in the yellow envelope that 50% percent of other participants were willing to at least 30 tasks for a \$15 donation and 25% were willing to at least 40 tasks for \$15 donation. Then what percent of previous experiment participants were actually willing to do more tasks than you?

- 95%
- 75%
- 50%
- 25%
- 0%

Screen 6: Summary

To sum up, the experiment proceeds as follows.

Step 1 You will choose how many tasks you are willing to do for each possible donation amount and you will guess what percent of people were willing to do more tasks than you for each

of your decisions. You will do this with sliders and multiple choice questions as shown below. *In Baseline:* In this session – as with 50% of all sessions – you *will not* learn any information on what others actually chose. (*In Treatment:* In this session – as with 50% of all sessions – you *will* learn any information on what others actually chose.)

Step 2 Each person will be randomly assigned a donation amount by drawing a ball out of a jar. A random number generator will then assign you a number of tasks. If this number of tasks is lower than the amount you were willing to do, then you will complete those tasks and we will donate to the A.D.R.P. on your behalf. The experiment will conclude after you finish your tasks (if any). We will hand you your show-up fee and completion bonus (\$15) as you exit the lab.

Instructions: Signal Extraction

Screen 1

Hello and welcome, today you will be participating in an experiment on economic decision making. Funds for this experiment have been provided by the University of California. You will be paid for your participation in this experiment. Your final payment will consist of a fixed \$5 show-up payment and a \$10 completion bonus. You will be paid privately in cash after the experiment has concluded. We ask that you please silence and put away all cell phones and any other personal electronic devices now and for the remainder of the experiment.

In this experiment you will be presented with an opportunity to complete a task for charity. For this project, we are partnering with Dr. James Rolfe – founder of the Afghan Dental Relief Project – to bring modern dental care to the very poor in Afghanistan. The A.D.R.P. is a charitable organization that provides free dental services and dental health education to the poorest families and individuals in the city of Kabul, the capital of Afghanistan. 100% of the

donations that you generate from this experiment will be used to purchase dental supplies and ship them to A.D.R.P.'s free dental clinic. Please take the time to carefully read the news article below that details the work Dr. Rolfe and the A.D.R.P. have accomplished and can continue to accomplish with your help.

Screen 2

Today, you will have the opportunity to complete tasks to generate donations to A.D.R.P.'s free dental clinic. The task will be to transcribe captchas. For each task, you will be shown an image of text. Your objective is to correctly type the text that is shown to you in the space provided. Only correctly transcribed captchas will be counted as completed tasks. You will have 3 chances per captcha. You will now be shown 5 captchas to transcribe so that you are familiar with the task. Please note that each captcha may consist of both capital and lower case letters, only the numbers 2-9, and special characters !, %, &, ?.

Screen 3

5 Sample Captchas

Screen 4: Tasks for Donations

For this experiment you will decide how many captchas you are willing to complete in order to generate a donation for the A.D.R.P.'s free dental clinic. That is, you will choose the maximum number of captchas you are willing to complete in exchange for a donation of \$20. You will be able to select between 0 and 50 tasks.

After you have made your decisions we will randomly assign each person a number of tasks. If the the number of tasks you are assigned is fewer than the maximum number you selected for yourself, then you will complete your assigned number of tasks and we will donate to A.D.R.P. on your behalf. If the randomly assigned number of tasks is higher than your chosen

number, then you will complete no tasks and no donation will be given to the A.D.R.P.

For example, suppose that you choose to do at most 25 tasks. Then if you are randomly assigned 20 tasks, then you will complete 20 tasks and \$20 will be given to the Afghan Dental Relief Project. On the other hand if you are randomly assigned 40 tasks, then you will do 0 tasks and no donation will be given to the A.D.R.P.

Each person will be assigned their own number of tasks. No one will learn your assigned number of tasks. Please note that the randomization device is “fair” – all possible task amounts will have equal probability of being assigned.

You will be free to leave the experiment as soon as you complete your tasks. You will be free to go immediately if you have no tasks to complete. Your completion bonus will be \$10 regardless of the number of tasks you have to do.

Identical Comprehension Questions for this section as the Anticipation Experiment

Screen 5: Seeing Other’s Choices

In addition to making your own choices, you will receive information about how your decisions compare to the decisions of the over 200 UCSD students who have already participated in this experiment. Before making your decisions, you will each draw an envelope at random. In each of the envelopes will be one of the two following statements on how many tasks previous participants were willing to complete in exchange for a \$20 donation to the A.D.R.P. The statements you could receive are:

- More than 50% of all previous participants were willing to do at least 20 tasks.
- Less than 25% of all previous participants were willing to do at least 20 tasks.

In Baseline: Only one of these statements is true. There are a total of 16 available envelopes at the front of the room. 8 of these envelopes contain the true statement and 8 envelopes contain the false statement. Therefore, when you pick an envelope, there will be a

50% chance that you will receive the true statement and a 50% chance that you will receive the false statement.

In Treatment: Only one of these statements is true. There are a total of 16 available envelopes at the front of the room. 15 of these envelopes contain the true statement and 1 envelope contains the false statement. Therefore, when you pick an envelope, there will be a 94% chance that you will receive the true statement and a 6% chance that you will receive the false statement.

Comprehension Questions, Set A

If 8-out-of-16 (*In Treatment:* 15-out-of-16) available envelopes contain the true statement, which of the following is most likely going to happen?

- I will receive the true statement.
- I will receive the false statement.
- I will receive the true or false statement with equal probability.

Suppose that you receive the statement that “more than 50% of all previous participants were willing to do at least 20 tasks.” What is the probability that this statement is TRUE and that *more than 50%* of all previous participants were willing to do at least 20 tasks?

- 94%
- 50%

Suppose that you receive the statement that “more than 50% of all previous participants were willing to do at least 20 tasks.” What is the probability that this statement is FALSE and that *less than 25%* of all previous participants were willing to do at least 20 tasks?

- 50%
- 6%

Comprehension Questions, Set B

If 8-out-of-16 (*in treatment*:15-out-of-16) available envelopes contain the true statement, what is the probability you will receive a true statement?

- 94%
- 50%

If 8-out-of-16 (*in treatment*:15-out-of-16) available envelopes contain the true statement, what is the probability you will receive a false statement?

- 50%
- 6%

If 8-out-of-16 (*in treatment*:15-out-of-16) available envelopes contain the true statement, what is the maximum number of true statements 12 participants could pick?

- 8
- 12

If 8-out-of-16 (*in treatment*:15-out-of-16) available envelopes contain the true statement, what is the maximum number of false statements 12 participants could pick?

- 8
- 1

We switched from comprehension question set B to A after the first few sessions of the signal extraction experiment. We believed set A would be more revealing of subject comprehension. The sessions with question sets A and B are balanced across treatments, and we found no difference in task choice dynamics between comprehension set A and B.

A.3 News Article about the A.D.R.P.

Afghan Dental Relief Project Ready for Next Level¹

By Kelsey Abkin

In 2003, Santa Barbara dentist James Rolfe came across an article about three women going to Afghanistan to treat victims of PTSD (post-traumatic stress disorder). Instead of putting the article down and going on with his day, Dr. Rolfe picked up the phone and asked to go along. The poverty and vulnerability he saw upon arriving in Afghanistan affected him immediately. Within a country of some 30 million people, Rolfe said he saw “resources in manpower and [natural] resources but no infrastructure to really use it.” Eleven years later, thanks to Rolfe and his Afghanistan Dental Relief Project, this is changing. The project started the Kabul Dental Clinic and Training Center, which offers basic dental services, and Rolfe is on the verge of expanding to a permanent dental clinic, hoping to extend nonbasic, often life-saving dental procedures.

The mission began when Rolfe returned to Afghanistan, this time with a homemade, portable dentists office and base camp. Before that, what passed for dental care in the war-torn country often amounted to a barber ripping out sore teeth without anesthetics. With 90 percent of Afghans having never seen a real dentist and 70 percent malnourished, dental problems were extreme. Abscesses were not uncommon and often led to septicemia, which can be lethal without antibiotics. Word that an American dentist had come offering free dental care spread fast among rural communities, and soon Rolfe was helping more than 60 people a day. Many of the orphaned boys whom Rolfe treated would become his assistants, thus leading to the Kabul School of Dental Technology.

After a car bomb, two scams, 100,000 patients treated, and 11 years since the birth of the

¹This is an abridged version of an article in the *Santa Barbara Independent* from 2014. This the the text that was shown to subjects. The full-length article can be found here: <https://www.independent.com/news/2014/jul/12/afghan-dental-relief-project-ready-next-level/>

Afghanistan Dental Relief Project, Rolfe is on the brink of taking it to a new level. He recently worked with Afghanistans Ministry of Public Health to obtain permits to provide more complex, nonbasic dental services to Afghans for a small fee, such as endodontic treatment or prosthetic restorations. For non-Afghans, however, the fee is equivalent to what they would pay in Dubai, and the treatment of only 16 non-Afghans covers the clinics entire monthly operating expenses.

Today, Rolfe can be found in his successful dentistry clinic near the Lobero Theatre. His self-built office surrounds his patients with the sounds of nature and artifacts reminiscent of a cultured life. He has managed to live a life performing dentistry for no cost in a Santa Barbara commune and now helping thousands in Afghanistan that intertwines his passion to help with his skills as a dentist. Rolfe, who is 75 years old, continues to work 115 hours a week and lives well below the poverty line. Except for his basic needs, he gives all he makes to his Afghanistan project, and hes nowhere near ready to slow down. In the future, he sees a first-rate dental infrastructure providing Afghans with health care, jobs and education. “We need to be more active,” Rolfe said. “If we feel something in our heart, we need to act on that, and that needs to form the basis of our existence.”

Appendix B

Procrastination in the Field

B.1 Additional Tables and Figures

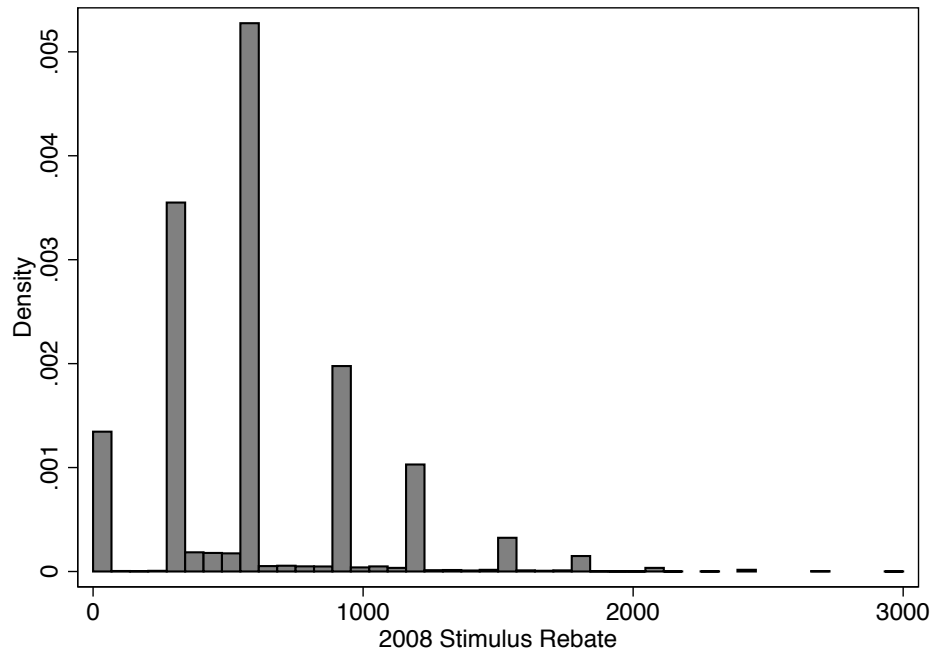


Figure B.1: Histogram of 2008 Stimulus Act Payments

Table B.1: Summary statistics

Variable	Roxbury Filers				Full	
	Experimental		Non-Participating		Obs	Mean
	Obs	Mean	Obs	Mean		
Age	1296	38.85 (14.90)	606	40.65 (16.22)	26228	37.80 (15.47)
AGI	1296	17452 (13991)	606	18498 (15179)	26248	16914 (13586)
Refund	1296	1317 (1571)	606	1205 (1482)	26248	1438 (1641)
Dependents	1296	0.505 (0.838)	606	0.418 (0.769)	26248	0.567 (0.90)
$1_{\text{Unemployed}}$	1296	0.101 (0.302)	606	0.112 (0.316)	26247	0.095 (0.29)
Female	1217	0.644 (0.479)	557	0.636 (0.482)	19881	0.646 (22.03)
Black	1211	0.761 (0.426)	556	0.781 (0.414)	19565	0.562 (0.49)
College	1180	0.138 (0.345)	534	0.133 (0.340)	18971	0.154 (0.36)
Days Until Deadline	1296	38.74 (21.28)	606	46.47 (24.00)	26248	43.07 (22.03)
$1_{\text{Direct Dep}}$	1296	0.444 (0.497)	606	0.398 (0.490)	26248	0.389 (0.49)

Table B.2: Aggregate Parameter Estimates 2007

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\delta}$	0.539 (0.004)	0.594 (0.001)	0.639 (0.000)	0.673 (0.000)	0.717 (0.000)	0.738 (0.000)
\hat{c}	3.279 (0.028)	5 -	10 -	20 -	50 -	75 -
# Observations	404191	404191	404191	404191	404191	404191
Log-Likelihood	-26323.63	-27521.71	-35519.37	-53546.81	-101781.32	-137074.84
$T - 1 : T - 7$ Average $\tilde{p}(1_t 1_t) - \hat{p}(1_t 1_t) :$	0.172	0.198	0.214	0.219	0.221	0.222
$T - 1 : T - 7$ Average $\tilde{q}(1_t 1_t) - \hat{q}(1_t 1_t) :$	0.0056	-0.0046	0.0057	0.0102	0.0111	0.0112

Notes: Structural estimates of exponential discounting, $\hat{\delta}$, and filing costs, \hat{c} , obtained via Maximum Likelihood Estimation for years 2005-2007. Columns (2) - (6) restrict costs to be between 5 and 75. Standard errors in parentheses. Also reported is the average excess conditional filing probability, $\tilde{p}(1_t|1_t) - \hat{p}(1_t|1_t)$, and the average excess unconditional filing probability, $\tilde{q}(1_t|1_t) - \hat{q}(1_t|1_t)$, over the seven days prior to the tax deadline. This table is a reproduction of Table 2.2 but only uses 2007 data.

Table B.3: Aggregate Parameter Estimates 2005 — 2007, $\mu = 0$

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\delta}$	0.534 (0.003)	0.600 (0.001)	0.642 (0.000)	0.675 (0.000)	0.719 (0.000)	0.739 (0.000)
\hat{c}	2.829 (0.016)	5 -	10 -	20 -	50 -	75 -
# Observations	1010387	1010387	1010387	1010387	1010387	1010387
Log-Likelihood	-65463.96	-70339.45	-91413.34	-135908.46	-253558.83	-339385.71
$T - 1 : T - 7$ Average						
$\bar{p}(1_t 1_t) - \hat{p}(1_t 1_t) :$	0.131	0.173	0.191	0.196	0.201	0.202
$T - 1 : T - 7$ Average						
$\bar{q}(1_t 1_t) - \hat{q}(1_t 1_t) :$	0.0096	-0.0043	0.0061	0.0104	0.0113	0.0116

Notes: Structural estimates of exponential discounting, $\hat{\delta}$, and filing costs, \hat{c} , obtained via Maximum Likelihood Estimation for years 2005-2007. Columns (2) - (6) restrict costs to be between 5 and 75. Standard errors in parentheses. Also reported is the average excess conditional filing probability, $\bar{p}(1_t|1_t) - \hat{p}(1_t|1_t)$, and the average excess unconditional filing probability, $\bar{q}(1_t|1_t) - \hat{q}(1_t|1_t)$, over the seven days prior to the tax deadline. This table is a reproduction of Table 2.2 with location parameter $\mu = 0$ (as opposed to $\mu = -\gamma$)

Table B.4: Aggregate Parameter Estimates 2005 — 2007, $k = 0$

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\delta}$	0.99999 (0.000004)	0.9999999 (0.000)	0.999999 (0.000)	0.999999 (0.000)	0.999999 (0.000)	0.999999 (0.000)
\hat{c}	2698.49 (1397.35)	5 -	10 -	20 -	50 -	75 -
# Observations	1010387	1010387	1010387	1010387	1010387	1010387
Log-Likelihood	-66180.702	-66182.362	-66182.361	-66182.359	-66182.352	-66182.346
$T - 1 : T - 7$ Average						
$\bar{p}(1_t 1_t) - \hat{p}(1_t 1_t) :$	-0.015	-0.018	-0.0189	-0.0189	-0.0190	-0.0190
$T - 1 : T - 7$ Average						
$\bar{q}(1_t 1_t) - \hat{q}(1_t 1_t) :$	-0.0078	-0.0025	-0.0025	-0.0024	-0.0025	-0.0025

Notes: Structural estimates of exponential discounting, $\hat{\delta}$, and filing costs, \hat{c} , obtained via Maximum Likelihood Estimation for years 2005-2007. Columns (2) - (6) restrict costs to be between 5 and 75. Standard errors in parentheses. Also reported is the average excess conditional filing probability, $\bar{p}(1_t|1_t) - \hat{p}(1_t|1_t)$, and the average excess unconditional filing probability, $\bar{q}(1_t|1_t) - \hat{q}(1_t|1_t)$, over the seven days prior to the tax deadline. This table is a reproduction of Table 2.2 with no delay in refund arrival ($k = 0$).

Table B.5: Aggregate Parameter Estimates 2005 — 2007, Paper Check

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\delta}$	0.631 (0.007)	0.701 (0.001)	0.743 (0.000)	0.770 (0.000)	0.803 (0.000)	0.818 (0.000)
\hat{c}	3.459 (0.030)	5 -	10 -	20 -	50 -	75 -
# Observations	623282	623282	623282	623282	623282	623282
Log-Likelihood	-40689.023	-41397.661	-47599.246	-63972.131	-110885.69	-145574.88
<i>T</i> - 1 : <i>T</i> - 7 Average						
$\bar{p}(1_t 1_t) - \hat{p}(1_t 1_t) :$	0.154	0.166	0.187	0.196	0.205	0.207
<i>T</i> - 1 : <i>T</i> - 7 Average						
$\bar{q}(1_t 1_t) - \hat{q}(1_t 1_t) :$	0.0044	-0.0095	-0.0015	0.0093	0.0131	.0135

Notes: Structural estimates of exponential discounting, $\hat{\delta}$, and filing costs, \hat{c} , obtained via Maximum Likelihood Estimation for years 2005-2007. Columns (2) - (6) restrict costs to be between 5 and 75. Standard errors in parentheses. Also reported is the average excess conditional filing probability, $\bar{p}(1_t|1_t) - \hat{p}(1_t|1_t)$, and the average excess unconditional filing probability, $\bar{q}(1_t|1_t) - \hat{q}(1_t|1_t)$, over the seven days prior to the tax deadline. This table is a reproduction of Table 2.2 with only those who receive their refunds via paper checks.

Table B.6: Aggregate Parameter Estimates 2005 — 2007, Direct Deposit

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\delta}$	0.515 (0.005)	0.588 (0.001)	0.634 (0.000)	0.669 (0.000)	0.714 (0.000)	0.734 (0.000)
\hat{c}	3.16 (0.033)	5 -	10 -	20 -	50 -	75 -
# Observations	387105	387105	387105	387105	387105	387105
Log-Likelihood	-24652.468	-25507.62	-30975.57	-44183.668	-81480.071	-109392.39
<i>T</i> - 1 : <i>T</i> - 7 Average						
$\bar{p}(1_t 1_t) - \hat{p}(1_t 1_t) :$	0.164	0.176	0.187	0.194	0.202	0.205
<i>T</i> - 1 : <i>T</i> - 7 Average						
$\bar{q}(1_t 1_t) - \hat{q}(1_t 1_t) :$	0.0040	-0.0047	-0.0003	0.0045	.0067	0.0071

Notes: Structural estimates of exponential discounting, $\hat{\delta}$, and filing costs, \hat{c} , obtained via Maximum Likelihood Estimation for years 2005-2007. Columns (2) - (6) restrict costs to be between 5 and 75. Standard errors in parentheses. Also reported is the average excess conditional filing probability, $\bar{p}(1_t|1_t) - \hat{p}(1_t|1_t)$, and the average excess unconditional filing probability, $\bar{q}(1_t|1_t) - \hat{q}(1_t|1_t)$, over the seven days prior to the tax deadline. This table is a reproduction of Table 2.2 with only those who receive their refunds via direct deposit.

Table B.7: Aggregate Parameter Estimates 2005 — 2007, Exclude Last 7 Days

	(1)	(2)	(3)	(4)	(5)	(6)
$\hat{\delta}$	0.527 (0.003)	0.590 (0.001)	0.638 (0.000)	0.673 (0.000)	0.717 (0.000)	0.738 (0.000)
\hat{c}	3.410 (0.018)	5 -	10 -	20 -	50 -	75 -
# Observations	916368	916368	916368	916368	916368	916368
Log-Likelihood	-62214.02	-64532.95	-82175.01	-122525.93	-230511.93	-309418.03
$T - 1 : T - 7$ Average						
$\bar{p}(1_t 1_t) - \hat{p}(1_t 1_t)$	0.131	0.173	0.191	0.196	0.201	0.202
$T - 1 : T - 7$ Average						
$\bar{q}(1_t 1_t) - \hat{q}(1_t 1_t)$	0.0096	-0.0043	0.0061	0.0104	0.0113	0.0116

Notes: Structural estimates of exponential discounting, $\hat{\delta}$, and filing costs, \hat{c} , obtained via Maximum Likelihood Estimation for years 2005-2007. Columns (2) - (6) restrict costs to be between 5 and 75. Standard errors in parentheses. Also reported is the average excess conditional filing probability, $\bar{p}(1_t|1_t) - \hat{p}(1_t|1_t)$, and the average excess unconditional filing probability, $\bar{q}(1_t|1_t) - \hat{q}(1_t|1_t)$, over the seven days prior to the tax deadline. This table is a reproduction of Table 2.2 while dropping the last 7 days prior to the deadline.

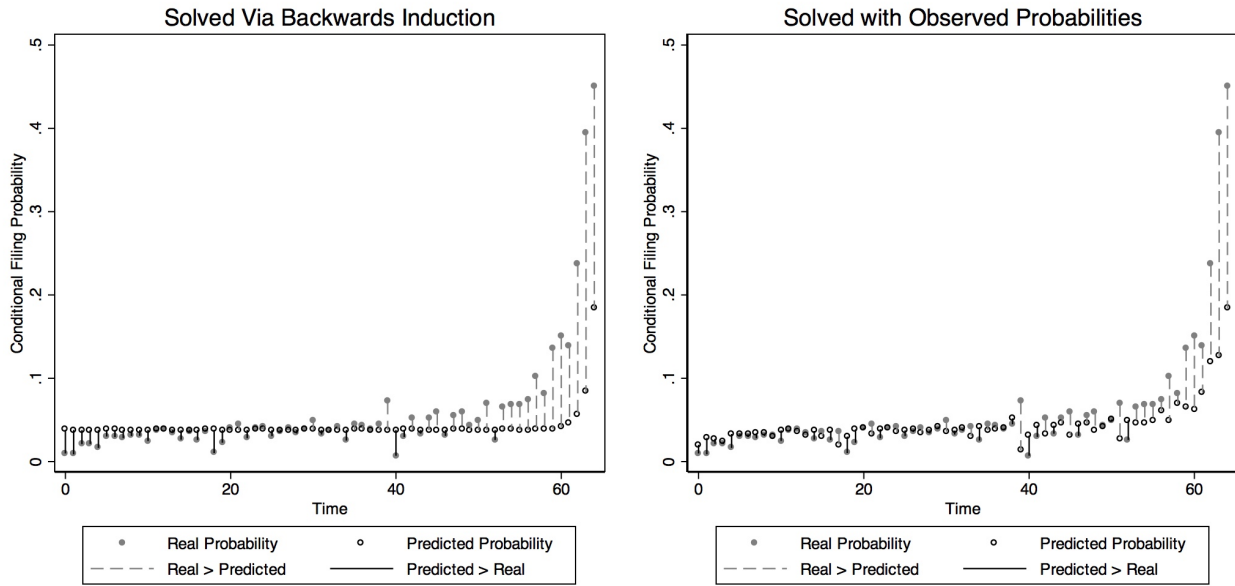


Figure B.2: Predicted Filing Times Using 1-Period-Ahead Observed Probabilities

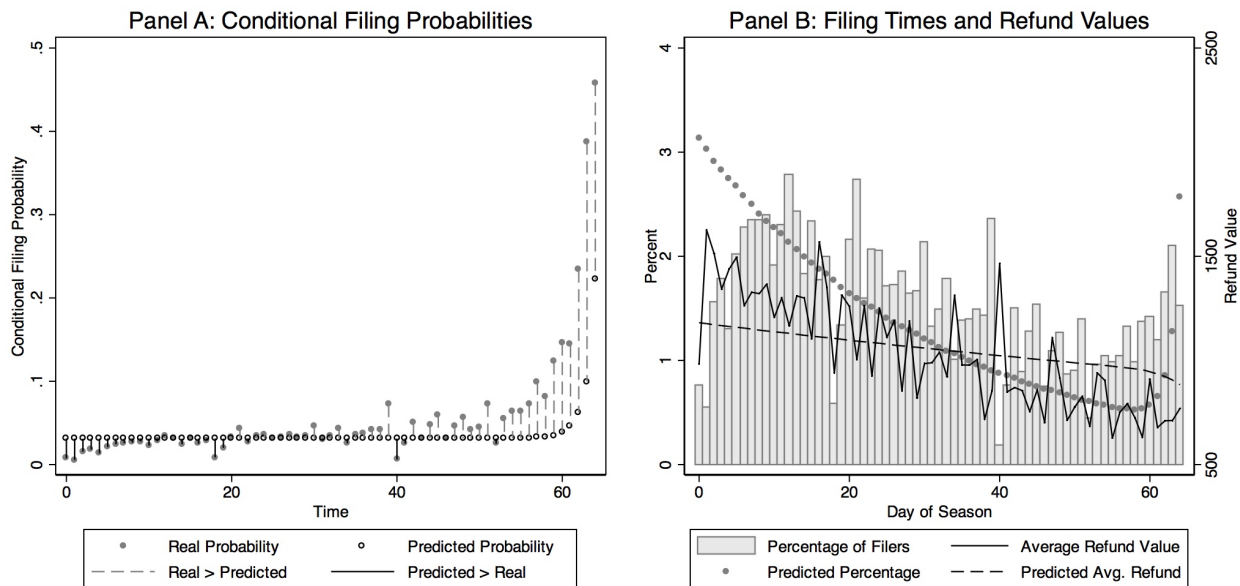


Figure B.3: Filing Times and Refund Values (Paper Check Only)

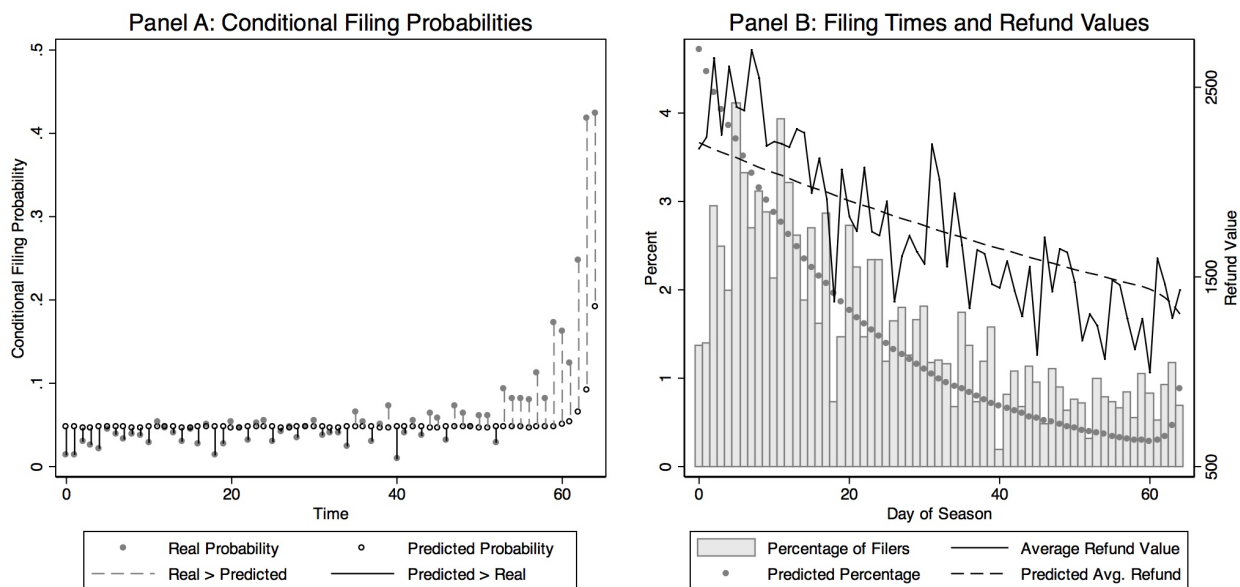


Figure B.4: Filing Times and Refund Values (Direct Deposit Only)

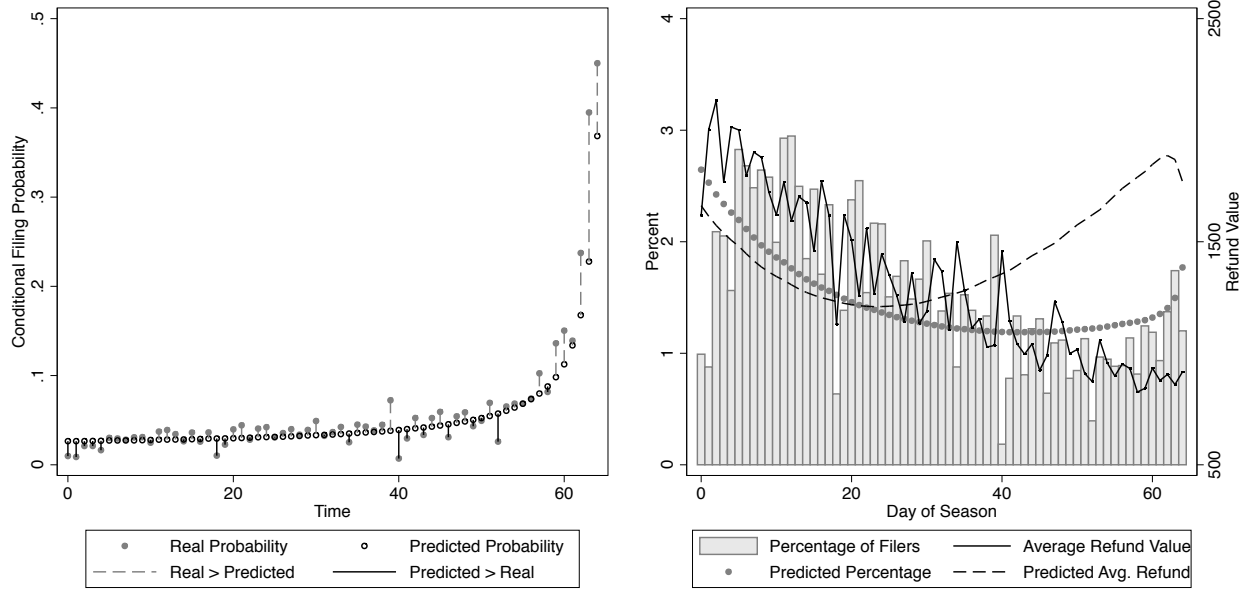


Figure B.5: Filing Times and Refund Values (δ Heterogeneous by Race)

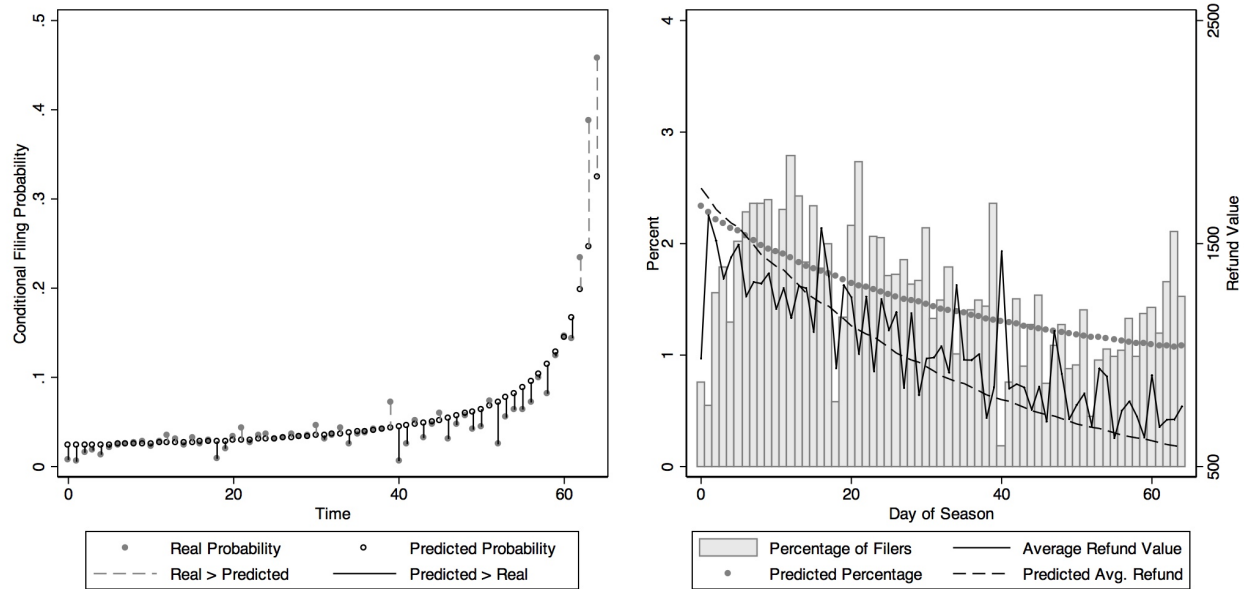


Figure B.6: Filing Times and Refund Values (Present Biased, Paper Check Only)

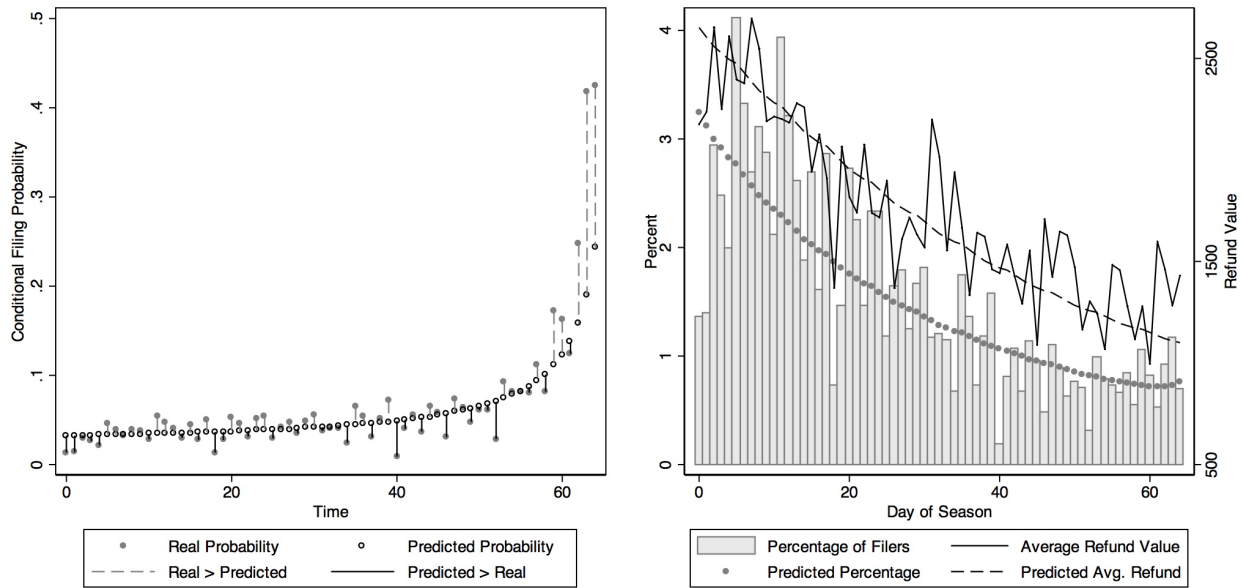


Figure B.7: Filing Times and Refund Values (Present Biased, Direct Deposit Only)

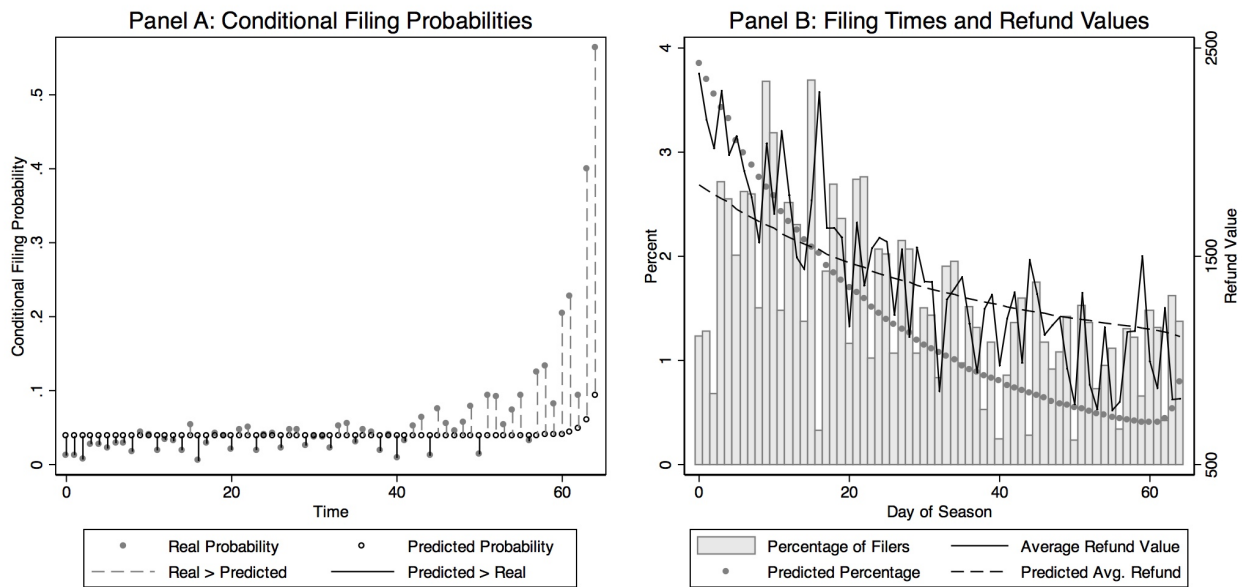


Figure B.8: Filing Times and Refund Values 2008 (Exponential)

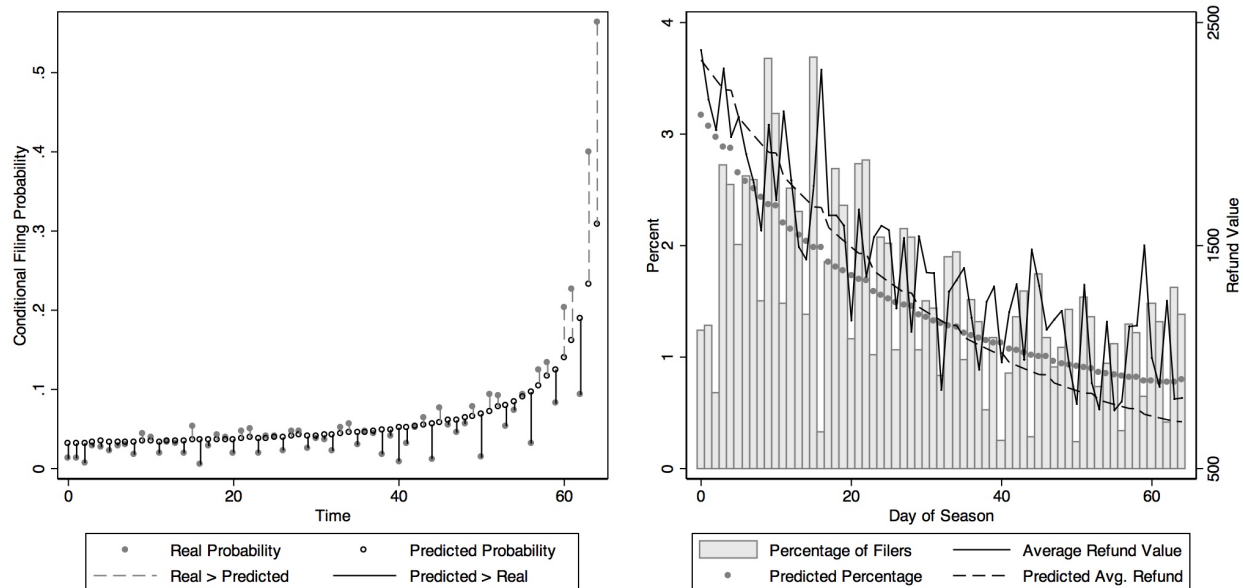


Figure B.9: Filing Times and Refund Values 2008 (Present Biased)

Table B.8: Difference-in-Difference Effect of Stimulus Payments (Placebo)

<i>Dep. Var.</i>	Linear Regression		Quantile Regression	
	(1)	(2)	(3)	(4)
$I_{\text{Stimulus Payment} > 0}$	-1.0031 (0.7062)	-0.6341 (0.7276)	-0.9228 (1.1911)	-0.6009 (1.1570)
$I_{\text{Year} = 2007}$	1.0079 (1.0067)	0.9760 (1.0075)	2.2275 (1.8180)	2.2980 (1.7438)
Interaction	-0.3734 (1.0505)	-0.5533 (1.0512)	-0.2586 (1.8638)	-0.4975 (1.7892)
Constant	31.2434*** (0.6427)	31.2478*** (0.6463)	28.1119*** (1.1218)	28.0560*** (1.0675)
Stimulus Determinants	No	Yes	No	Yes
# Observations	15972	15972	15972	15972

Notes: This table demonstrates the relationship between when one filed (Day Completed) and whether one received a 2008 Stimulus Payment. This is placebo test of the interaction between $I_{\text{Year} = 2007}$ and $I_{\text{Stimulus Payment} > 0}$. Standard errors are clustered by individual.

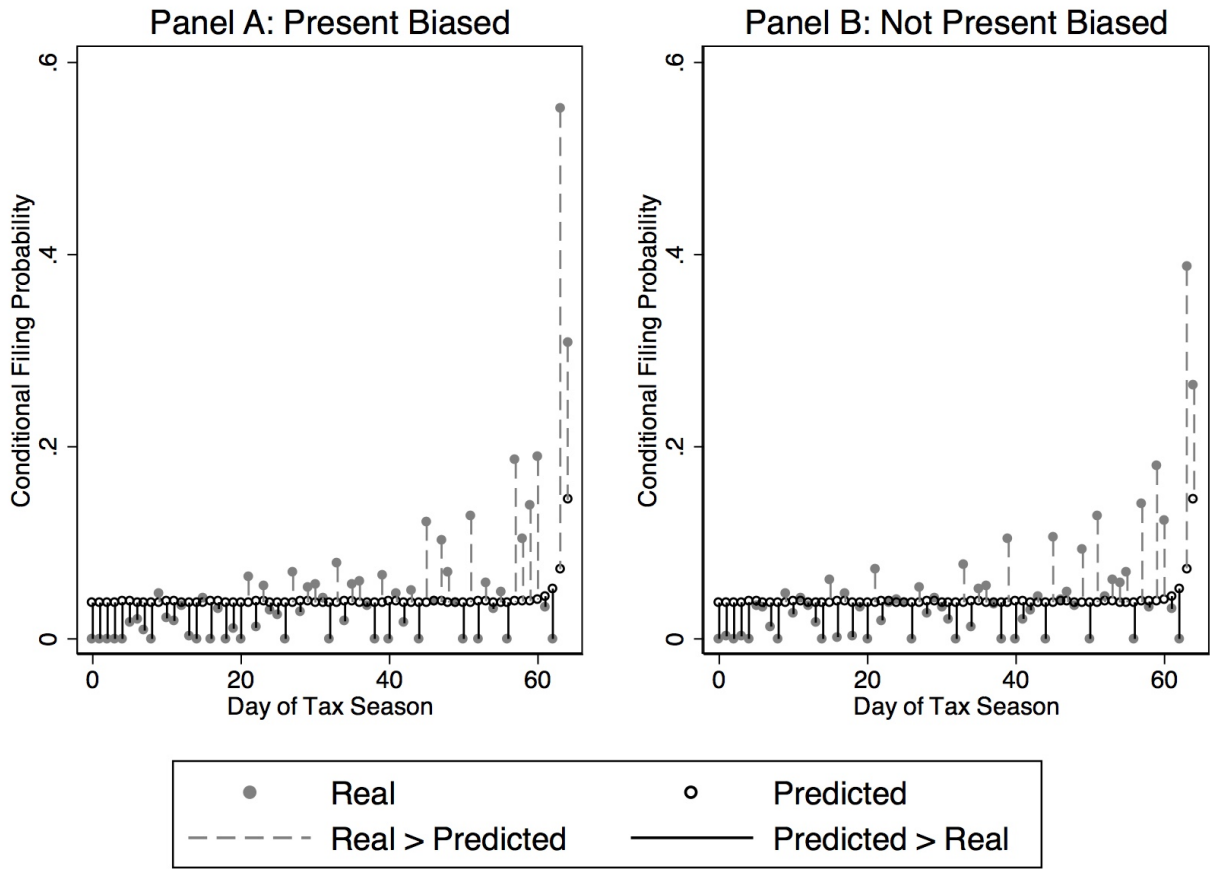


Figure B.10: Predicted and Actual Filing Behavior by Present Bias (Un-smoothed)

B.2 Tax Filing as Dynamic Discrete Choice

We detail the methodology we implement which follows from Hotz and Miller (1993) and Arcidiciano and Ellickson (2011).

Recall that there are N individual tax filers, indexed by i . Time is discrete, indexed by t , with T denoting the period of the tax deadline. In each period, tax filers take actions a_{it} . They either decide to postpone filing ($a_{it} = 0$) or to file ($a_{it} = 1$). Let f_{it} denote the individual's filing status in period t such that $f_{it} = 1$ if the individual has not yet filed by period $t - 1$ and $f_{it} = 0$ if the individual has filed by period $t - 1$. We assume that each individual will receive a positive refund, b_i , constant through time and known to the researcher and the filer. This refund is to be received in a fixed number of periods, k , after filing. The state variables known to the researcher are $x_{it} = (f_{it}, b_i)$, which is Markov.

We assume that the cost of filing consists of a fixed component c and time-varying idiosyncratic shocks ε_{it} . These shocks are contemporaneously observed by the filer but unobserved to the researcher. These shocks may depend on the choice of filing and hence we write $\varepsilon(a_{it})$. We assume $\varepsilon(a_{it})$ is independent and identically distributed over time.

Filer utility is additively separable. The utility of filing in period t is

$$\delta^k b_{it} - c + \varepsilon(1)$$

when $a_{it} = 1$ and $f_{it} = 1$. The variable δ^k is a k period *exponential* discount factor homogeneous in the population of filers. Utility is $\varepsilon(0)$ if $a_{it} = 0$ and $f_{it} = 1$. As such, the flow utility can be written

$$u(x_{it}, a_{it}) + \varepsilon_{it}(a_{it}) = a_{it} f_{it} (\delta^k b_i - c) + \varepsilon_{it}(a_{it}).$$

With these flow utilities, the filer maximizes the present discounted value of filing-related utilities by choosing α_i^* , a set of decision rules for all possible realizations of observed and unobserved

state variables in each time period. That is,

$$\alpha_i^* = \arg \max_{\alpha_i} E_{\alpha_i} \sum_{t=1}^T \delta^{t-1} [u(x_{it}, a_{it}) + \varepsilon_{it}(a_{it})].$$

The corresponding value function at time t can be defined recursively as

$$V_{it}(x_{it}, \varepsilon_{it}) = \max_{a_{it}} [u(x_{it}, a_{it}) + \varepsilon_{it}(a_{it}) + \delta E [V_{i,t+1}(x_{i,t+1}, \varepsilon_{i,t+1}) | x_{it}, a_{it}]]$$

We define the *ex-ante* value function, $\bar{V}_{it}(x_{it})$, obtained by integrating over the possible realizations of shocks as

$$\bar{V}_{it}(x_{it}) \equiv \int V_{it}(x_{it}, \varepsilon_{it}) g(\varepsilon_{it}) d\varepsilon_{it}.$$

We additionally define the conditional value function $v_{it}(x_{it}, a_{it})$ as the present discounted value (net of the shocks ε_{it}) of choosing a_{it} and behaving optimally from period $t + 1$ on as

$$v_{it}(x_{it}, a_{it}) \equiv u(x_{it}, a_{it}) + \delta \int \bar{V}_{i,t+1}(x_{i,t+1}) f(x_{i,t+1} | x_{it}, a_{it}) dx_{i,t+1}.$$

The optimal decision rule at time t solves

$$\alpha_{it}(x_{it}, \varepsilon_{it}) = \arg \max_{a_{it}} v_{it}(x_{it}, a_{it}) + \varepsilon_{it}(a_{it}).$$

B.2.1 Type-1 Extreme Value Errors

Following the logic of static discrete choice problems, the probability of observing an action a_{it} conditional on x_{it} is found by integrating out ε_{it} from the optimal decision rule.

$$\begin{aligned} p(a_{it} | x_{it}) &= \int \mathbf{1}[\alpha_{it}(x_{it}, \varepsilon_{it}) = a_{it}] g(\varepsilon_{it}) d\varepsilon_{it} \\ &= \int \mathbf{1}[\arg \max_{a_{it}} v_{it}(x_{it}, a_{it}) + \varepsilon_{it}(a_{it}) = a_{it}] g(\varepsilon_{it}) d\varepsilon_{it} \end{aligned}$$

Hence, if we are able to form the conditional value function $v_{it}(x_{it}, a_{it})$, standard methods can be applied.

In order to obtain choice probabilities and other constructs in closed form, we assume a type-1 extreme value distribution, fixing the location parameter equal to minus Euler's constant, $-\gamma = -0.5772$, and the scale parameter to 1.¹ This leads to a dynamic logit model where the probability of an arbitrary choice a_{it} is given by

$$p(a_{it}|x_{it}) = \frac{\exp(v_{it}(x_{it}, a_{it}))}{\sum_{a'_{it}} \exp(v_{it}(x_{it}, a'_{it}))}.$$

Or,

$$p(a_{it}|x_{it}) = \frac{1}{\sum_{a'_{it}} \exp(v_{it}(x_{it}, a'_{it}) - v_{it}(x_{it}, a_{it}))}.$$

As in the standard logit, the probability of any action being taken is expressed in terms of relative utility values or utility differences. Here, however, the relevant utility values are the conditional value functions. The conditional value functions carry with them both the current flow payoffs and discounted considerations of taking the prescribed action and then acting optimally from then on.

Importantly, under the above distributional assumption we also have the ex-ante value

¹The cumulative distribution function of a type-1 extreme value random variable, X , $Prob(X \leq x) = F(x; \mu, \lambda) = e^{-e^{-(x-\mu)/\lambda}}$, is summarized by location parameter, μ , and scale parameter, λ . The expectation is

$$E[X] = \mu + \gamma\lambda,$$

such that fixing location $\mu = -\gamma$ and $\lambda = 1$ ensures the shocks are mean zero. In most situations, imposing mean zero shocks is inconsequential as choices are driven by the *difference* between action-specific shocks. In period T , however, the individual must file if she has not yet, and so in period $T - 1$ the individual forecasts only one relevant shock in the subsequent period. Additionally, in all prior periods, choosing to file in the period stops the problem and sets all future flow utilities to zero, while choosing not to file exposes the decisionmaker to future shocks. Under the traditional assumption that $\mu = 0$, such future shocks would yield additional option value to not filing in any period. In Appendix Table B.3 we re-estimate the specifications of Table 2.2 with the traditional assumption of $\mu = 0$ and show that estimated discount factors and costs are both slightly reduced under this assumption. Note as well, that the variance of a type-1 extreme value random variable is $\lambda \frac{\pi^2}{6}$, such that restricting $\lambda = 1$ also restricts the variance of the shocks to be $\frac{\pi^2}{6}$. In section ?? we analyze behavior with alternate assumptions for λ , and ensure that both our estimated model and simulations with $\mu = -\gamma$ and $\lambda = 1$ predict identical behavior.

function in closed form:

$$\bar{V}_{it}(x_{it}) = \ln \left[\sum_{a'_{it}} \exp(v_{it}(x_{it}, a'_{it})) \right].$$

A powerful observation by Hotz and Miller (1993) recognizes that

$$\bar{V}_{it}(x_{it}) = -\ln[p(a_{it}^*|x_{it})] + v_{it}(x_{it}, a_{it}^*)$$

for some arbitrary action taken at time t , a_{it}^* . This expresses the ex-ante value of being at a given state as the conditional value of taking an arbitrary action adjusted for a penalty that the arbitrary action might not be optimal. We can substitute this in to our equation for the conditional value function to obtain

$$v_{it}(x_{it}, a_{it}) = u(x_{it}, a_{it}) + \delta \int (v_{i,t+1}(x_{i,t+1}, a_{i,t+1}^*) - \ln[p(a_{i,t+1}^*|x_{i,t+1})]) f(x_{i,t+1}|x_{it}, a_{it}) dx_{i,t+1}$$

The components of the conditional value function are contemporaneous flow payoffs, $u(x_{it}, a_{it})$, the one period ahead conditional value function for the arbitrary action, $v_{i,t+1}(x_{i,t+1}, a_{i,t+1}^*)$, conditional choice probabilities for the arbitrary action $p(a_{i,t+1}^*|x_{i,t+1})$, and state transition probabilities, $f(x_{i,t+1}|x_{it}, a_{it})$. These constructs are obtainable in the following ways:

Formulating Contemporaneous Flow Payoffs: The flow payoffs are established as $\delta^k b_i - c$ if a person enters period t without having filed and files in that period. The flow payoffs are 0 otherwise. The refund value noted in Table 2.1, provides the value b_i . This refund will be received in k periods with the assumptions that for direct deposit filers, $k = 14$ while for paper check filers $k = 21$.² The parameters to estimate are the discount factor, δ , and filing costs, c .

Outside of the estimated discount factor and filing costs, the contemporaneous flow payoffs are driven by refund values and direct deposit status. Our estimation strategy takes these

²We follow the IRS refund cycle charts for 2005-2008 to arrive at these values of k .

values as exogenous and does not model the choice of refund value or payment method. We do not have access to a potential instrument for direct deposit choice (e.g., variation in local bank services) or refund values (e.g., variation in refundability of tax credits). As such, we provide separate estimates for individuals with and without direct deposit and find qualitatively similar results (see Tables B.5 and B.6 for further detail). For refund values we provide out-of-sample examination of the 2008 Stimulus Payments which provided plausibly exogenous variation in refund values and timing. Failure to account for endogenous refund choice in estimation should lead to substantial mis-prediction out-of-sample, while we show the predictions closely match behavior for our preferred specification (see section 2.4.2 for details).

Obtaining Conditional Choice Probabilities: We wish to have an estimated probability for a_{it} given the state vector x_{it} . Our states are the benefit amount, b_i , and whether someone has not already filed, f_{it} . These can be calculated with simple bin estimators.

$$\tilde{p}(a_{it}|x_{it}) = \frac{\sum_{i=1}^N \mathbf{1}(a_{it} = a_t, x_{it} = x_t)}{\sum_{i=1}^N \mathbf{1}(x_{it} = x_t)}$$

Obtaining State Transition Probabilities: Our only states are the benefit amounts b_i and the filing status f_{it} . The benefit amount is unchanging through time, and conditional upon f_{it} and the choice a_{it} , $f_{i,t+1}$ can be known with certainty. Hence, all the state transition probabilities are 1.

Obtaining Arbitrary Action Payoff from Terminating Actions: Our setup is such that there exist terminating actions. Once a filer files, no further choice can be made. The decision problem is no longer dynamic. This is important because if we think of this terminating action of filing as the arbitrary action, $a_{i,t+1}^*$, then the remaining analysis is dramatically simplified. The terminating action makes all future payoffs zero and makes future shocks irrelevant. Filling in $a_{i,t+1}^* = 1$ as

the arbitrary action, we know that the $t + 1$ conditional value function will be

$$v_{i,t+1}(1, b_i, 1) = \delta_i^k b_i - c$$

if $f_{i,t+1} = 1$. Otherwise, the individual has already filed, $f_{i,t+1} = 0$, and this along with all future values are deterministically zero.

B.2.2 Likelihood Formulation

Our primary equation for the value of a given action given a particular state is

$$v_{it}(x_{it}, a_{it}) = u(x_{it}, a_{it}) + \delta \int (v_{i,t+1}(x_{i,t+1}, a_{i,t+1}^*) - \ln[p(a_{i,t+1}^* | x_{i,t+1})]) f(x_{i,t+1} | x_{it}, a_{it}) dx_{i,t+1}$$

The critical case for our estimator is $x_{it} = (f_{it}, b_i) = (1, b_i)$. The individual has not yet filed their taxes and decides between filing today and not filing today. Filing today yields immediate costs and discounted benefits. It also transitions the future filing state to $f_{i,t+1} = 0$, such that all future flow payoffs and the future ex-ante value function is zero. Together these yield

$$v_{it}(1, b_i, 1) = \delta^k b_i - c + \delta \int 0$$

$$v_{it}(1, b_i, 1) = \delta^k b_i - c$$

Now, consider the individual who chooses to not file. Filing today yields zero costs and zero benefits. It advances time, but the state in the future will be $x_{i,t+1} = (1, b_{i,t+1})$ with probability 1. Given this state and the arbitrary action that the individual files, the value of this option is simply calculated as well.

$$v_{it}(x_{it}, a_{it}) = u(x_{it}, a_{it}) + \delta \int (v_{i,t+1}(x_{i,t+1}, 1) - \ln[p(1 | x_{i,t+1})]) f(x_{i,t+1} | x_{it}, a_{it}) dx_{i,t+1}$$

$$v_{it}(1, b_i, 0) = 0 + \delta[\delta^k b_i - c] - \delta \ln[p(1|x_{i,t+1})]$$

$$v_{it}(1, b_i, 1) = \delta^{k+1} b_i - \delta c - \delta \ln[p(1|x_{i,t+1})]$$

We can evaluate the difference between these two conditional value functions as:

$$\begin{aligned} v_{it}(1, b_i, 0) - v_{it}(1, b_i, 1) &= \\ \left[\delta^{k+1} b_i - \delta c - \delta \ln[p(1|x_{i,t+1})] \right] - \left[\delta^k b_i - c \right] &= \\ (\delta^{k+1} - \delta^k) b_i - (\delta - 1) c - \delta \ln(p(1|1, b_i)). \end{aligned}$$

Under the error distribution, we have:

$$p(a_{it}|x_{it}) = \frac{1}{\sum_{a_{it}'} \exp(v_{it}(x_{it}, a_{it}') - v_{it}(x_{it}, a_{it}))}.$$

Or,

$$p(1_{it}|1_{it}, b_i) = \frac{1}{1 + \exp \left[(\delta^{k+1} - \delta^k) b_i - (\delta - 1) c - \delta \ln(p(1_{i,t+1}|1_{i,t+1}, b_i)) \right]}.$$

This represents the likelihood contribution of observation t for individual i given the decision maker has not filed yet. Note that we only need to consider those periods up until the time when the person files. Once they file, the utility consequences of filing are eliminated and the likelihood contribution is zero for such observations. Let D_i be the filing date of a given individual. The grand log likelihood is written as

$$\mathcal{L} = \sum_{i=1}^N \left[\sum_{t=1}^{D_i} \ln[p(1_{it}|1_{it}, b_i)] \right]$$

B.3 Refund Arrival Delay and Identification

B.3.1 Starting and Limiting Values

We start with the following conditional filing probabilities.

$$p_{T-t-1} = \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_{T-t})]}$$

$$p_{T-1} = \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c]}.$$

Now, suppose that $0 < p_{T-t-1} < p_{T-t} \leq 1$ then:

$$\begin{aligned} p_{T-t-2} &= \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_{T-t-1})]} \\ &< \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_{T-t})]} = p_{T-t-1} \end{aligned}$$

Therefore, we know $p_{T-t-2} < p_{T-t-1}$ for all t – i.e. p_{T-t} is monotonically decreasing in t since t may equal 1. Further suppose $0 < p_{T-t} \leq 1$ and $0 < \delta < 1$, then

$$0 < \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_{T-t})]} = p_{T-t-1} \forall t$$

.

Which tells us the sequence is bounded below by 0. Therefore, we may solve for the limit:

$$p_L = \frac{1}{1 + \exp [(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_L)]}$$

$$\iff 1 = p_L + \exp [(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_L)] p_L$$

$$\Leftrightarrow \ln\left(\frac{1-p_L}{p_L}\right) = (\delta^{k+1} - \delta^k)b_i - (\delta - 1)c - \delta \ln(p_L)$$

$$\Leftrightarrow (1-p_L)p_L^{\delta-1} = \exp[(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c]$$

But this means that

$$p_{T-1}^{-1} - 1 = \exp\left[(\delta^{k+1} - \delta^k)b_i - (\delta - 1)c\right]$$

which implies

$$(1-p_L)p_L^{\delta-1} = p_{T-1}^{-1} - 1$$

B.3.2 Single Coincidence

Let p_t represent the conditional probability of filing in period t under parameters (δ, c) and \tilde{p}_t represent the conditional probability of filing in period t under parameters $(\tilde{\delta}, \tilde{c})$. WLOG let $\delta > \tilde{\delta}$. We will show that either $p_{T-1} \leq \tilde{p}_{T-1}$ and $p_{T-t} < \tilde{p}_{T-t}$ for all $t > 1$, or that $p_{T-1} > \tilde{p}_{T-1}$ and that if there exists some \hat{t} for which $p_{T-\hat{t}} < \tilde{p}_{T-\hat{t}}$ then there either exists t' such that $p_{T-t} < \tilde{p}_{T-t}$ for all $t > t'$ and $p_{T-t} > \tilde{p}_{T-t}$ for all $0 < t < t'$.

Case 1: Suppose $p_{T-1} \leq \tilde{p}_{T-1}$. Then we know that

$$\frac{1}{1 + e^{(\delta-1)(\delta^k b - c)}} \leq \frac{1}{1 + e^{(\tilde{\delta}-1)(\tilde{\delta}^k b - \tilde{c})}}$$

$$\Leftrightarrow (\delta - 1)(\delta^k b - c) \geq (\tilde{\delta} - 1)(\tilde{\delta}^k b - \tilde{c})$$

Now consider any p_{T-t} and \tilde{p}_{T-t} such that $0 < p_{T-t} \leq \tilde{p}_{T-t} < 1$. But $\delta > \tilde{\delta}$, therefore we know that

$$(\delta - 1)(\delta^k b - c) - \delta \ln(p_{T-t}) > (\tilde{\delta} - 1)(\tilde{\delta}^k b - \tilde{c}) - \tilde{\delta} \ln(\tilde{p}_{T-t})$$

$$\iff \frac{1}{1 + \exp((\delta - 1)(\delta^k b - c) - \delta \ln(p_{T-t}))} < \frac{1}{1 + \exp((\tilde{\delta} - 1)(\tilde{\delta}^k b - \tilde{c}) - \tilde{\delta} \ln(\tilde{p}_{T-t}))}$$

So we know that $p_{T-t-1} < \tilde{p}_{T-t-1}$. But t was arbitrary and $p_{T-1} \leq \tilde{p}_{T-1}$ by assumption. Therefore, $p_{T-t} < \tilde{p}_{T-t}$ for all t .

Case 2: Suppose that $p_{T-1} > \tilde{p}_{T-1}$ and that $p_{T-t} > \tilde{p}_{T-t}$ for all $0 < t < t'$. Let $p_{T-t'} \leq \tilde{p}_{T-t'}$. Then we know that:

$$\frac{1}{1 + \exp((\delta - 1)(\delta^k b - c) - \delta \ln(p_{T-t'+1}))} \leq \frac{1}{1 + \exp((\tilde{\delta} - 1)(\tilde{\delta}^k b - \tilde{c}) - \tilde{\delta} \ln(\tilde{p}_{T-t'+1}))}$$

$$\iff (\delta - 1)(\delta^k b - c) - \delta \ln(p_{T-t'+1}) \geq (\tilde{\delta} - 1)(\tilde{\delta}^k b - \tilde{c}) - \tilde{\delta} \ln(\tilde{p}_{T-t'+1})$$

$$\iff (\delta - 1)(\delta^k b - c) - (\tilde{\delta} - 1)(\tilde{\delta}^k b - \tilde{c}) \geq \delta \ln(p_{T-t'+1}) - \tilde{\delta} \ln(\tilde{p}_{T-t'+1})$$

Now suppose that $p_{T-t'-n} < \tilde{p}_{T-t'-n}$ for some $n > 0$. Then we know that $p_{T-t'-n} - \tilde{p}_{T-t'-n} < p_{T-t'+1} - \tilde{p}_{T-t'+1}$. Further, we know $p_{T-t'-n} < p_{T-t'+1}$ and $\tilde{p}_{T-t'-n} < \tilde{p}_{T-t'+1}$.³

³This comes from the fact that if $0 < p_{T-t-1} < p_{T-t} \leq 1$ then $p_{T-t-2} < p_{T-t-1}$ (shown above in B.3.1) and

Therefore, since $\delta > \tilde{\delta}$ we know that:

$$\delta \ln(p_{T-t'+1}) - \tilde{\delta} \ln(\tilde{p}_{T-t'+1}) > \delta \ln(p_{T-t'-n}) - \tilde{\delta} \ln(\tilde{p}_{T-t'-n})$$

$$\implies (\delta - 1)(\delta^k b - c) - (\tilde{\delta} - 1)(\tilde{\delta}^k b - \tilde{c}) > \delta \ln(p_{T-t'-n}) - \tilde{\delta} \ln(\tilde{p}_{T-t'-n})$$

$$\implies p_{T-t'-n-1} < \tilde{p}_{T-t'-n-1}$$

But n was arbitrary, so we know $p_{T-t} < \tilde{p}_{T-t}$ for all $t > t'$ and $p_{T-t} > \tilde{p}_{T-t}$ for all $0 < t < t'$.

B.3.3 Level Curve Simulations

We first address the importance in identification of $k > 0$. Recall that the likelihood expression of observation t for individual i is:

$$p(1_{it}|1_{it}, b_i) = \frac{1}{1 + \exp[\delta^k(\delta - 1)b - (\delta - 1)c - \delta \ln(p(1_{i,t+1}|1_{i,t+1}, b_i))]}.$$

If we let $k = 0$ then the likelihood expression becomes:

$$p(1_{it}|1_{it}, b_i) = \frac{1}{1 + \exp[(\delta - 1)b - (\delta - 1)c - \delta \ln(p(1_{i,t+1}|1_{i,t+1}, b_i))]}.$$

Note that when $k = 0$, b and c must be weighted equally in the likelihood expression. This presents a problem in identification. Since $(\delta - 1)b - (\delta - 1)c = (\delta - 1)(b - c)$, the level sets of $p(1_{it}|1_{it}, b_i)$ in the (δ, c) plane coincide across time periods when solving $p(1_{it}|1_{it}, b_i)$ recursively. This is illustrated in Figure B.11. Figure B.11 displays the level sets of $p(1_{it}|1_{it}, b_i)$ — the fact that $p_{T-2} < p_{T-1}$ and $\tilde{p}_{T-2} < \tilde{p}_{T-1}$ necessarily.

assuming $\delta = 0.535$, $c = 10$, refund = 50 – over time periods $t = 61, 62, 63$, and 64 in the (δ, c) plane. We observe that the levels sets separate over different time periods when $k = 14$, but find no such separation when $k = 0$.

In our identification, we take $p(1_{it}|1_{it}, b_i)$, $p(1_{i,t+1}|1_{i,t+1}, b_i)$, and b as given, and we estimate δ and c via maximum likelihood. We know by observation that $p(1_{it}|1_{it}, b_i) \in (0, 1]$, $\ln(p(1_{i,t+1}|1_{i,t+1}, b_i)) \in [-6.59, 0]$ and that the mean refund is \$1419. As such, for any $0 < \delta < 1$ and reasonable cost parameter c , $\exp[(\delta - 1)b - (\delta - 1)c - \delta \ln(p(1_{i,t+1}|1_{i,t+1}, b_i))] \approx 0 \implies p(1_{it}|1_{it}, b_i) = \frac{1}{1 + \exp[(\delta - 1)b - (\delta - 1)c - \delta \ln(p(1_{i,t+1}|1_{i,t+1}, b_i))]} \approx 1$. That is, for reasonable values of c , $p(1_{it}|1_{it}, b_i)$ is computationally 1. However, if $k > 0$ the relative size of the refund does not necessitate that $p(1_{it}|1_{it}, b_i) = 1$. This point is illustrated in Figure B.12. Figure B.12 recreates figure B.11 but assumes refund = 1400. We observe that the levels sets separate over different time periods when $k = 14$. However, when $k = 0$, $p(1_{it}|1_{it}, b_i) = 1$ for all c when fixing any δ . As such, the level sets not only coincide over all time periods, but coincide over (δ, c) within a time period. Unsurprisingly, when running our estimator under the assumption $k = 0$ we find a flat likelihood over all values of c – see Figure B.4. In Figure B.13 we see further separation of level curves between time periods when increasing k to 21. Hence, differential direct deposit use, which generates differences in k , delivers additional identifying variation.

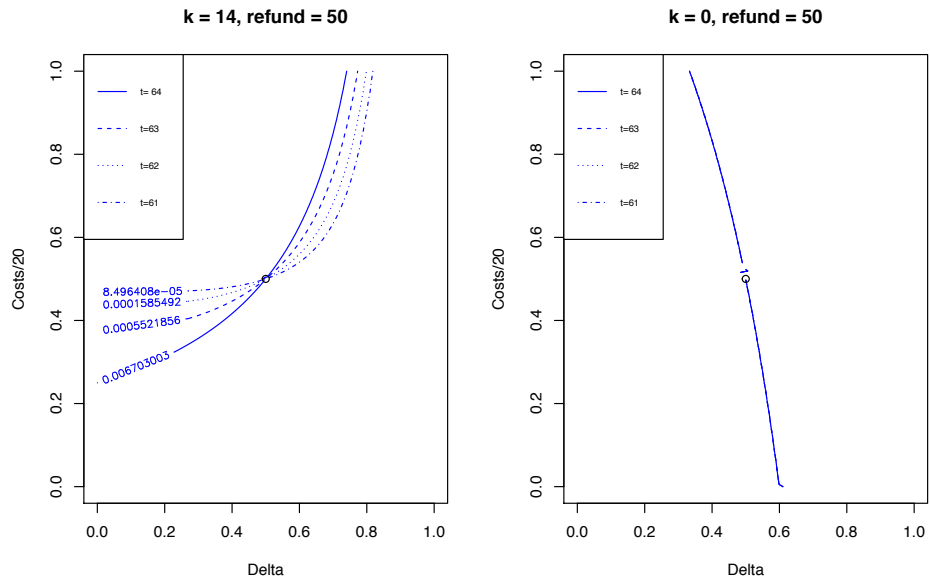


Figure B.11: Conditional Probability Level Curves assuming $\delta = 0.535, c = 10$

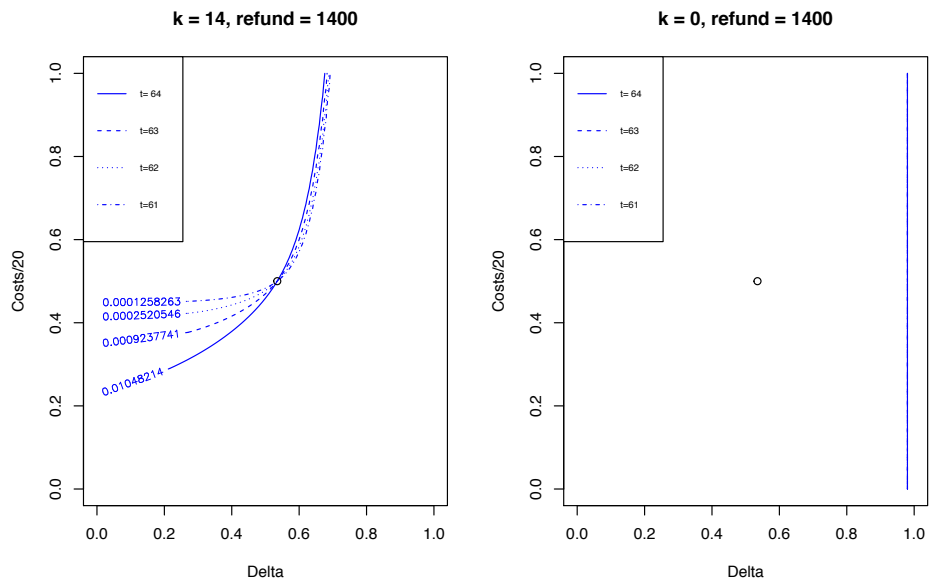


Figure B.12: Conditional Probability Level Curves assuming $\delta = 0.535, c = 10$

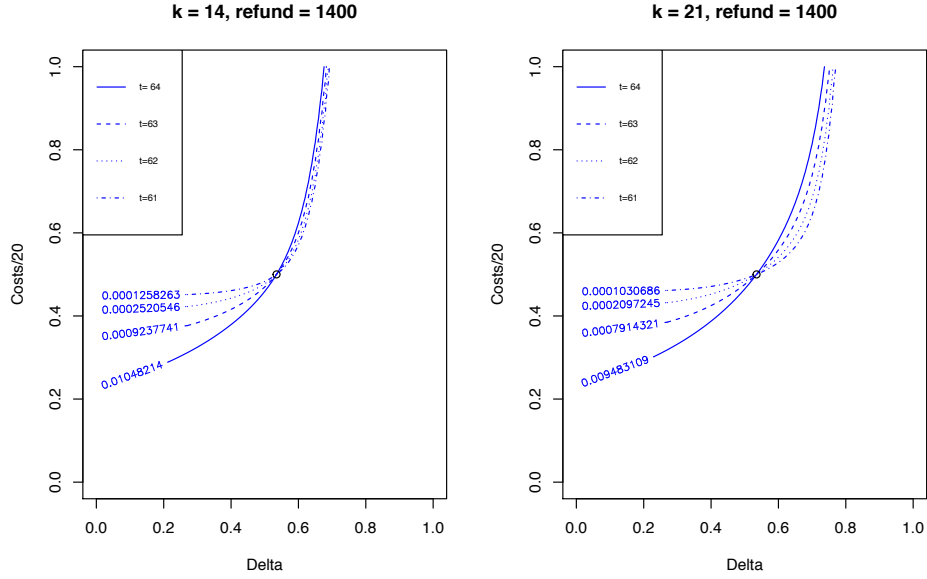


Figure B.13: Conditional Probability Level Curves assuming $\delta = 0.535, c = 10$

B.4 Naive $\beta - \delta$ derivation

Recall the formulation for the exponential agent:

$$V_{it}(x_{it}, \varepsilon_{it}) = \max_{a_{it}} [(\delta^k b - c)a_{it} + \varepsilon_{it}(a_{it}) + \delta E [V_{i,t+1}(x_{i,t+1}, \varepsilon_{i,t+1}) | x_{it}, a_{it}]]$$

Since, the naive present-biased agent believes he will act exponentially in all future periods we may write his current value function as:

$$W_{it}(x_{it}, \varepsilon_{it}) = \max_{a_{it}} [(\beta \delta^k b - c)a_{it} + \varepsilon_{it}(a_{it}) + \beta \delta E [V_{i,t+1}(x_{i,t+1}, \varepsilon_{i,t+1}) | x_{it}, a_{it}]],$$

where the expected future value function $V_{i,t+1}(x_{i,t+1}, \varepsilon_{i,t+1})$ is that of the exponential agent. Define the conditional naive present-biased value function, $w_{it}(x_{it}, a_{it})$, as the present discounted value (net of the shocks ε_{it}) of choosing a_{it} and behaving optimally (under the belief of exponen-

tial) from period $t + 1$ on as

$$w_{it}(x_{it}, a_{it}) \equiv (\beta \delta^k b - c)a_{it} + \beta \delta \int \bar{V}_{i,t+1}(x_{i,t+1}) f(x_{i,t+1} | x_{it}, a_{it}) dx_{i,t+1},$$

where $\bar{V}_{i,t+1}(x_{i,t+1})$, is

$$\bar{V}_{i,t+1}(x_{i,t+1}) \equiv \int V_{i,t+1}(x_{i,t+1}, \varepsilon_{i,t+1}) g(\varepsilon_{i,t+1}) d\varepsilon_{i,t+1}.$$

The optimal decision rule at time t solves

$$\alpha_{it}(x_{it}, \varepsilon_{it}) = \arg \max_{a_{it}} w_{it}(x_{it}, a_{it}) + \varepsilon_{it}(a_{it}).$$

Under the assumed Type-1 extreme value distribution, the probability of an arbitrary choice a_{it} is given by

$$p_{\beta \delta}(a_{it} | x_{it}) = \frac{1}{\sum_{a'_{it}} \exp(w_{it}(x_{it}, a'_{it}) - w_{it}(x_{it}, a_{it}))}.$$

As before we have the exponential ex-ante value function in closed form.

$$\begin{aligned} \bar{V}_{it}(x_{it}) &= \ln \left[\sum_{a'_{it}} \exp(v_{it}(x_{it}, a'_{it})) \right]. \\ \bar{V}_{it}(x_{it}) &= -\ln[p_n(a_{it}^* | x_{it})] + v_{it}(x_{it}, a_{it}^*), \end{aligned}$$

for some arbitrary action taken at time t , a_{it}^* . This expresses the ex-ante value of being at a given state as the conditional value of taking an arbitrary action adjusted for a penalty that the arbitrary action might not be optimal, now under the naive belief of exponential behavior, p_n . We can substitute this in to our equation for the conditional value function to obtain

$$w_{it}(x_{it}, a_{it}) = (\beta \delta^k b - c)a_{it} + \beta \delta \int (v_{i,t+1}(x_{i,t+1}, a_{i,t+1}^*) - \ln[p_n(a_{i,t+1}^* | x_{i,t+1})]) f(x_{i,t+1} | x_{it}, a_{it}) dx_{i,t+1}$$

Recognizing that

$$w_{it}(1, b_i, 1) = \beta \delta^k b_i - c$$

$$w_{it}(1, b_i, 0) = \beta \delta^{k+1} b_i - \beta \delta c - \beta \delta \ln[p_n(1|x_{i,t+1})],$$

we obtain

$$p_{\beta \delta}(1_{it}|1_{it}, b_i) = \frac{1}{1 + \exp [\beta (\delta^{k+1} - \delta^k) b - (\beta \delta - 1) c - \beta \delta \ln(p_n(1_{i,t+1}|1_{i,t+1}, b_i))]}.$$

In effect, $p_n(1_{i,t+1}|1_{i,t+1}, b_i)$ constructed with a given δ and c replaces the rational expectations bin estimates of future filing probabilities. The remaining parameter to estimate is β in a likelihood formulation that is altogether similar to that provided for an exponential agent.

B.5 IRS Announcement of H.R. 5140



Economic Stimulus Payment Notice

Dear Taxpayer:

We are pleased to inform you that the United States Congress passed and President George W. Bush signed into law the Economic Stimulus Act of 2008, which provides for economic stimulus payments to be made to over 130 million American households. Under this new law, you may be entitled to a payment of up to \$600 (\$1,200 if filing a joint return), plus additional amounts for each qualifying child.

We are sending this notice to let you know that based on this new law the IRS will begin sending the one-time payments starting in May. To receive a payment in 2008, individuals who qualify will not have to do anything more than file a 2007 tax return. The IRS will determine eligibility, figure the amount, and send the payment. This payment should not be confused with any 2007 income tax refund that is owed to you by the federal government. Income tax refunds for 2007 will be made separately from this one-time payment.

For individuals who normally do not have to file a tax return, the new law provides for payments to individuals who have a total of \$3,000 or more in earned income, Social Security benefits, and/or certain veterans' payments. Those individuals should file a tax return for 2007 to receive a payment in 2008.

Individuals who qualify may receive as much as \$600 (\$1,200 if married filing jointly). Even if you pay no income tax but have a total of \$3,000 or more in earned income, Social Security benefits, and/or certain veterans' payments, you may receive a payment of \$300 (\$600 if married filing jointly).

In addition, individuals eligible for payments may also receive an additional amount of \$300 for each child qualifying for the child tax credit.

For taxpayers with adjusted gross income (AGI) of more than \$75,000 (or more than \$150,000 if married filing jointly), the payment will be reduced or phased out completely.

To qualify for the payment, an individual, spouse, and any qualifying child must have a valid Social Security number. In addition, individuals cannot receive a payment if they can be claimed as a dependent of another taxpayer or they filed a 2007 Form 1040NR, 1040NR-EZ, 1040-PR, or 1040-SS.

All individuals receiving payments will receive a notice and additional information shortly before the payment is made. In the meantime, for additional information, please visit the IRS website at www.irs.gov.

B.6 Excerpts from Technical Explanation of H.R. 5140

B.6.1 Explanation of Provision

B.6.2 In general

The provision includes a recovery rebate credit for 2008 which is refundable. The credit mechanism (and the issuance of checks described below) is intended to deliver an expedited fiscal stimulus to the economy.

The credit is computed with two components in the following manner.

B.6.3 Basic credit

Eligible individuals receive a basic credit (for the first taxable year beginning) in 2008 equal to the greater of the following:

- Net income tax liability not to exceed 600(1,200 in the case of a joint return).
- 300(600 in the case of a joint return) if: (1) the eligible individual has qualifying income of at least 3,000; *or* (2) *the eligible individual has a net income tax liability of at least 1* and gross income greater than the sum of the applicable basic standard deduction amount and one personal exemption (two personal exemptions for a joint return).

An eligible individual is any individual other than: (1) a nonresident alien; (2) an estate or trust; or (3) a dependent. For these purposes, “net income tax liability” means the excess of the sum of the individual’s regular tax liability and alternative minimum tax over the sum of all nonrefundable credits (other than the child credit). Net income tax liability as determined for these purposes is not reduced by the credit added by this provision or any credit which is refundable under present law. Qualifying income is the sum of the eligible individual’s: (a) earned income; (b) social security benefits (within the meaning of sec. 86(d)); and (c) veteran’s

payments (under Chapters 11, 13, or 15 of title 38 of the U. S. Code). The definition of earned income has the same meaning as used in the earned income credit except that it includes certain combat pay and does not include net earnings from self-employment which are not taken into account in computing taxable income.

B.6.4 Qualifying child credit

If an individual is eligible for any amount of the basic credit the individual also may be eligible for a qualifying child credit. The qualifying child credit equals \$300 for each qualifying child of such individual. For these purposes, the child credit definition of qualifying child applies.

B.6.5 Limitation based on gross income

The amount of the credit (i.e., the sum of the amounts of the basic credit and the qualifying child credit) is phased out at a rate of five percent of adjusted gross income above certain income levels. The beginning point of this phase-out range is \$75,000 of adjusted gross income (\$150,000 in the case of joint returns).

Examples of rebate determination

Example 1. - The amount of the credit (i.e., the sum of the amounts of the basic credit and the qualifying child credit) is phased out at a rate of five percent of adjusted gross income above certain income levels. The beginning point of this phase-out range is 75,000 of adjusted gross income (150,000 in the case of joint returns).

Example 6. - A married taxpayer filing jointly has \$175,000 in earned income, two qualifying children, and a net tax liability of \$31,189 (the taxpayer's actual liability after the child credit also is \$31,189 as the joint income is too high to qualify). The taxpayer meets the qualifying income test and the net tax liability test. The taxpayer will, in the absence of the rebate phase-out provision, receive a rebate of \$1,800, comprising \$1,200 (greater of \$600 or

net tax liability not to exceed \$1,200), and \$300 per child. The phase-out provision reduces the total rebate amount by five percent of the amount by which the taxpayer's adjusted gross income exceeds \$150,000. Five percent of \$25,000 (\$175,000 minus \$150,000) equals \$1,250. The taxpayer's rebate is thus \$1,800 minus \$1,250, or \$550.

B.7 Instructions



2008 BOSTON EITC CAMPAIGN RAFFLE QUESTIONS

The following questions are asked for research purposes only. We will never share your personal information with any organization or its representatives. Please note that any winnings may be taxable.

- Use a No. 2 pencil only.
- Do not use ink, ballpoint or felt tip pens.
- Make solid marks that fill the oval completely
- Erase cleanly any marks you wish to change
- Make no stray marks on this form
- Do not fold, tear or mutilate this form

As a tax filer at this Volunteer Income Tax Assistance site you are automatically entered in a raffle in which you could win up to \$50. Just follow the directions below:

How It Works:
In the boxes below you are asked to choose between smaller payments closer to today and larger payments further in the future. For each row, choose one payment: either the smaller, sooner payment or the later, larger payment. When you return this completed form, you will receive a raffle ticket. If you are a winner, the raffle ticket will have a number on it from 1 to 22. These numbers correspond to the numbered choices below. You will be paid your chosen payment. The choices you make could mean a difference in payment of more than \$35, so ... CHOOSE CAREFULLY!!!

RED BLOCK (Numbers 1 through 7) : Decide between payment today and payment in one month
BLACK BLOCK (Numbers 8 through 15): Decide between payment today and payment in six months
BLUE BLOCK (Numbers 16 through 22): Decide between payment in six months and payment in seven months.

Rules and Eligibility:
For each possible number below, state whether you would like the earlier, smaller payment or the later, larger payment. Only completed raffle forms are eligible for the raffle.
All prizes will be sent to you by normal mail and will be paid by money order. One out of ten raffle tickets will be a winner. You can obtain your raffle ticket as soon as your tax filing is complete. You may not participate in the raffle if you are associated with the EITC campaign (volunteer, business associate, etc.) or an employee (or relative of an employee) of the Federal Reserve Bank of Boston or the Federal Reserve System.

TODAY VS. ONE MONTH FROM TODAY

WHAT WILL YOU DO IF YOU GET A NUMBER BETWEEN 1 AND 7?

Decide for **each** possible number if you would like the smaller payment for sure **today** or the larger payment for sure in **one month**? Please answer for each possible number (1) through (7) by filling in one box for each possible number.

*Example: If you prefer \$49 today in Question 1 mark as follows: \$49 today or \$50 in one month
 If you prefer \$50 in one month in Question 1, mark as follows: \$49 today or \$50 in one month*

If you get number (1):	Would you like to receive <input type="checkbox"/> \$49 today	or <input type="checkbox"/> \$50 in one month
If you get number (2):	Would you like to receive <input type="checkbox"/> \$47 today	or <input type="checkbox"/> \$50 in one month
If you get number (3):	Would you like to receive <input type="checkbox"/> \$44 today	or <input type="checkbox"/> \$50 in one month
If you get number (4):	Would you like to receive <input type="checkbox"/> \$40 today	or <input type="checkbox"/> \$50 in one month
If you get number (5):	Would you like to receive <input type="checkbox"/> \$35 today	or <input type="checkbox"/> \$50 in one month
If you get number (6):	Would you like to receive <input type="checkbox"/> \$29 today	or <input type="checkbox"/> \$50 in one month
If you get number (7):	Would you like to receive <input type="checkbox"/> \$22 today	or <input type="checkbox"/> \$50 in one month

TODAY VS. SIX MONTHS FROM TODAY

WHAT WILL YOU DO IF YOU GET A NUMBER BETWEEN 8 AND 15?

Now, decide for **each** possible number if you would like the smaller payment for sure **today** or the larger payment for sure in **six months**? Please answer each possible number (8) through (15) by filling in one box for each possible number.

If you get number (8): Would you like to receive \$49 **today** or \$50 in **six months**

If you get number (9): Would you like to receive \$47 **today** or \$50 in **six months**

If you get number (10): Would you like to receive \$44 **today** or \$50 in **six months**

If you get number (11): Would you like to receive \$40 **today** or \$50 in **six months**

If you get number (12): Would you like to receive \$35 **today** or \$50 in **six months**

If you get number (13): Would you like to receive \$29 **today** or \$50 in **six months**

If you get number (14): Would you like to receive \$22 **today** or \$50 in **six months**

If you get number (15): Would you like to receive \$14 **today** or \$50 in **six months**

SIX MONTHS FROM TODAY VS. SEVEN MONTHS FROM TODAY

WHAT WILL YOU DO IF YOU GET A NUMBER BETWEEN 16 AND 22?

Decide for **each** possible number if you would like the smaller payment for sure **in six months** or the larger payment for sure **in seven months**? Please answer for each possible number (16) through (22) by filling in one box for each possible number.

If you get number (16): Would you like to receive \$49 **in six months** or \$50 **in seven months**

If you get number (17): Would you like to receive \$47 **in six months** or \$50 **in seven months**

If you get number (18): Would you like to receive \$44 **in six months** or \$50 **in seven months**

If you get number (19): Would you like to receive \$40 **in six months** or \$50 **in seven months**

If you get number (20): Would you like to receive \$35 **in six months** or \$50 **in seven months**

If you get number (21): Would you like to receive \$29 **in six months** or \$50 **in seven months**

If you get number (22): Would you like to receive \$22 **in six months** or \$50 **in seven months**

References

- Afriat, Sidney N. 1967. "The construction of utility functions from expenditure data." *International economic review*, 8(1): 67–77.
- Aguirregabiria, Victor, and Pedro Mira. 2007. "Sequential estimation of dynamic discrete games." *Econometrica*, 75(1): 1–53.
- Akerlof, George A. 1980. "A theory of social custom, of which unemployment may be one consequence." *The quarterly journal of economics*, 94(4): 749–775.
- Akerlof, George A, and Rachel E Kranton. 2000. "Economics and Identity*." *The Quarterly Journal of Economics*, 115(3): 715–753.
- Allen, Eric J., Patricia M. Dechow, Devin G. Pope, and George Wu. Forthcoming. "Reference Dependent Preferences: Evidence from Marathon Runners." *Management Science*.
- Alpizar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman. 2008. "Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica." *Journal of Public Economics*, 92(5): 1047–1060.
- Andreoni, James, and Charles Sprenger. 2012. "Estimating Time Preferences with Convex Budgets." *American Economic Review*.
- Andreoni, James, and Douglas B. Bernheim. 2009. "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica*, 77(5): 1607–1636.
- Arcidiciano, Peter, and Paul B. Ellickson. 2011. "Practical Methods for Estimation of Dynamic Discrete Choice Models." *Annual Review of Economics*, 3: 363–394.
- Ariely, Dan, and Klaus Wertenbroch. 2002. "Procrastination, Deadlines, and Performance: Self-Control by Precommitment." *Psychological Science*, 13(3): 219–224.
- Ashraf, Nava, Dean Karlan, and Wesley Yin. 2006. "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines." *Quarterly Journal of Economics*,

121(1): 635–672.

- Augenblick, Ned, and Matthew Rabin. 2016. “An Experiment on Time Preference and Misprediction in Unpleasant Tasks.” *Working Paper*.
- Augenblick, Ned, Muriel Niederle, and Charles Sprenger. 2015. “Working Over Time: Dynamic Inconsistency in Real Effort Tasks.” *Quarterly Journal of Economics*, 130(3): 1067–1115.
- Bajari, Patrick, C Lanier Benkard, and Jonathan Levin. 2007. “Estimating dynamic models of imperfect competition.” *Econometrica*, 75(5): 1331–1370.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2010. “Social Incentives in the Workplace.” *The Review of Economic Studies*, 77(2): 417–458.
- Banerjee, Abhijit V. 1992. “A simple model of herd behavior.” *The Quarterly Journal of Economics*, 107(3): 797–817.
- Bénabou, Roland, and Jean Tirole. 2006. “Incentives and Prosocial Behavior.” *The American Economic Review*, 96(5): 1652–1678.
- Benabou, Roland, and Jean Tirole. 2011. “Laws and norms.” National Bureau of Economic Research.
- Benjamin, Daniel, James Choi, and Joshua Strickland. 2010. “Social Identity and Preferences.” *American Economic Review*, 100(4).
- Benzarti, Youssef. 2015. “How Taxing is Filing? Leaving Money on the Table Because of Hassle Costs.” *Working Paper*.
- Bernheim, B. Douglas. 1994. “A Theory of Conformity.” *Journal of Political Economy*, 102(5): 841–877.
- Bernheim, B Douglas, and Christine L Exley. 2015. “Understanding conformity: an experimental investigation.” *Working Paper*.
- Birkhoff, Garrett. 1946. “Tres observaciones sobre el álgebra lineal.” *Univ. Nac. Tucumán Rev. Ser. A*, 5: 147–151.
- Bisin, Alberto, and Kyle Hyndman. 2014. “Present-Bias, Procrastination and Deadlines in a Field Experiment.” *Working Paper*.
- Brown, Jeffrey R., and Alessandro Previtro. 2014. “Procrastination, Present-Biased Preferences, and Financial Behaviors.” *Working Paper*.

- Brown, Jeffrey R., Anne M. Farrell, and Scott J. Weisbenner. 2016. "Decision-making approaches and the propensity to default: Evidence and Implications." *Journal of Financial Economics*, 121: 477–495.
- Burks, Stephan, Jeffrey Carpenter, Lorenz Goette, and Aldo Rustichini. 2009. "Cognitive skills affect economic preferences, strategic behavior, and job attachment." *Proceedings of the National Academy of Science*, 106(19): 7745–7750.
- Bursztyn, Leonardo, and Robert Jensen. 2017. "Social Image and Economic Behavior in the Field: Identifying, Understanding and Shaping Social Pressure." *Annual Review of Economics*, 9.
- Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman. 2014. "Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions." *Econometrica*, 82(4): 1273–1301.
- Bursztyn, Leonardo, Georgy Egorov, and Robert Jensen. 2017. "Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education." *The Review of Economic Studies*.
- Carvalho, Leandro, Stephan Meier, and Stephanie Wang. 2016. "Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday." *American Economic Review*, 106(2): 260–284.
- Chabris, Christopher F., David Laibson, and Jonathon P. Schuldt. 2008. "Intertemporal Choice." In *The New Palgrave Dictionary of Economics*, ed. Steven N. Durlauf and Larry Blume. London: Palgrave Macmillan.
- Chambers, Christopher P., Federico Echenique, and Eran Shmaya. 2014. "The axiomatic structure of empirical content." *American Economic Review*, 104(8): 2303–2319.
- Chen, Yan, F Maxwell Harper, Joseph Konstan, and Sherry Xin Li. 2010. "Social comparisons and contributions to online communities: A field experiment on movielens." *The American economic review*, 100(4): 1358–1398.
- Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *The Quarterly Journal of Economics*, 126(2): 749–804.
- Coller, Maribeth, and Melonie B. Williams. 1999. "Eliciting individual discount rates." *Experimental Economics*, 2: 107–127.
- Cubitt, Robin P., and Daniel Read. 2007. "Can Intertemporal Choice Experiments Elicit Preferences for Consumption." *Experimental Economics*, 10(4): 369–389.

- Dahl, Gordon B, Katrine V Løken, and Magne Mogstad. 2014. "Peer Effects in Program Participation." *American Economic Review*, 104(7): 2049–2074.
- Dana, Jason, Roberto A. Weber, and Jason Xi Kuang. 2007. "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." *Economic Theory*, 33(1).
- Dean, Mark, and Anja Sautmann. 2016. "Credit Constraints and the Measurement of Time Preferences." *Working Paper*.
- DellaVigna, Stefano, and Ulrike Malmendier. 2006. "Paying Not to Go to the Gym." *American Economic Review*, 96(3): 694–719.
- DellaVigna, Stefano, Attila Lindner, Balazs Reizer, and Johannes F. Schmieder. Forthcoming. "Reference-Dependent Job Search: Evidence from Hungary." *Quarterly Journal of Economics*.
- Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2006. "Dynamic inconsistency predicts self-control problems in humans." *Working Paper*.
- Dufló, Esther, and Emmanuel Saez. 2002. "Participation and investment decisions in a retirement plan: the influence of colleagues' choices." *Journal of Public Economics*, 85(1): 121–148.
- Echenique, Federico, and Kota Saito. 2015. "Savage in the market." *Econometrica*, 83(4): 1467–1495.
- Efron, Bradley, and Robert J. Tibshirani. 1994. *An Introduction to the Bootstrap*. Chapman and Hall.
- Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2016. "Bunching at the kink: implications for spending responses to health insurance contracts." *Journal of Public Economics*.
- Epstein, Larry G. 2000. "Are probabilities used in markets?" *Journal of Economic Theory*, 91(1): 86–90.
- Exley, Christine L. 2016. "Excusing Selfishness in Charitable Giving: The Role of Risk." *The Review of Economic Studies*, 83(2): 587–628.
- Falk, Armin, and Andrea Ichino. 2006. "Clean evidence on peer effects." *Journal of Labor Economics*, 24(1): 39–57.
- Fang, Hamming, and Dan Silverman. 2009. "Time-inconsistency and welfare program participation: Evidence from the NLSY." *International Economic Review*, 116(2): 1043–1077.
- Fang, Hanming, and Yang Wang. 2015. "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions." *International*

- Economic Review*, 56(2): 565–596.
- Festinger, Leon. 1954. “A Theory of Social Comparison Processes.” *Human relations*, 7(2): 117–140.
- Fishburn, Peter C. 1975. “Separation theorems and expected utilities.” *Journal of Economic Theory*, 11(1): 16–34.
- Foster, Andrew D, and Mark R Rosenzweig. 1995. “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture.” *Journal of Political Economy*, 103(6): 1176–1209.
- Frakes, Michael D., and Melissa F. Wasserman. 2016. “Procrastination in the Workplace: Evidence from the U.S. Patent Office.” *NBER Working Paper No. 22987*.
- Frederick, Shane, George Loewenstein, and Ted O’Donoghue. 2002. “Time discounting and time preference: A critical review.” *Journal of Economic Literature*, 40(2): 351–401.
- Frey, Bruno S, and Stephan Meier. 2004. “Social comparisons and pro-social behavior: Testing” conditional cooperation” in a field experiment.” *The American Economic Review*, 94(5): 1717–1722.
- Gine, Xavier, Dean Karlan, and Jonathan Zinman. 2010. “Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation.” *American Economic Journal - Applied Economics*, 2(4): 213–235.
- Goette, Lorenz, David Huffman, and Stephan Meier. 2006. “The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups.” *The American Economic Review*, 96(2): 212–216.
- Green, Richard C., and Sanjay Srivastava. 1986. “Expected utility maximization and demand behavior.” *Journal of Economic Theory*, 38(2): 313–323.
- Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams. 2002. “Estimating individual discount rates in Denmark: A field experiment.” *American Economic Review*, 92(5): 1606–1617.
- Harrison, Glenn W., Morten I. Lau, Elisabet E. Rutstrom, and Melonie B. Williams. 2005. “Eliciting risk and time preferences using field experiments: Some methodological issues.” In *Field experiments in economics*. Vol. 10 (Research in Experimental Economics), ed. Jeffrey Carpenter, Glenn W. Harrison and John A. List. Greenwich and London: JAI Press.
- Heffetz, Ori, Ted O’Donoghue, and Henry S. Schneider. 2016. “Forgetting and Heterogeneity in Task Delay: Evidence from New York City Parking-Ticket Recipients.” *Working Paper*.

- Holt, Charles A., and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects." *The American Economic Review*, 92(5): 1644–1655.
- Hotz, V. Joseph, and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies*, 60(3): 497–529.
- Hoxby, Caroline. 2000. "Peer effects in the classroom: Learning from gender and race variation." National Bureau of Economic Research.
- Ifcher, John, and Homa Zarghamee. 2011. "Happiness and Myopia: Evidence from a Random-Assignment Experiment." *American Economic Review*, 101(7): 3109–3129.
- Jacobson, Sarah, and Ragan Petrie. 2009. "Learning from Mistakes: What Do Inconsistent Choices Over Risk Tell Us?" *Journal of Risk and Uncertainty*, 38(2).
- Jones, Stephen R G. 1984. *The economics of conformism*. Blackwell.
- Katz, Michael L, and Carl Shapiro. 1986. "Technology adoption in the presence of network externalities." *Journal of political economy*, 94(4): 822–841.
- Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan. 2010. "Self-Control and the Development of Work Arrangements." *American Economic Review, Papers and Proceedings*, 100(2): 624–628.
- Kennan, John, and James R Walker. 2011. "The effect of expected income on individual migration decisions." *Econometrica*, 79(1): 211–251.
- Koopmans, Tjalling C. 1960. "Stationary Ordinal Utility and Impatience." *Econometrica*, 28(2): 287–309.
- Krupka, Erin, and Roberto a Weber. 2009. "The focusing and informational effects of normas on pro-social behavior." *Journal of Economic Psychology*, 30(3): 307–320.
- Krupka, Erin L, and Roberto A Weber. 2013. "Identifying social norms using coordination games: Why does dictator game sharing vary?" *Journal of the European Economic Association*, 11(3): 495–524.
- Kubler, Felix, Larry Selden, and Xiao Wei. 2014. "Asset demand based tests of expected utility maximization." *American Economic Review*, 104(11): 3459–3480.
- Laibson, David. 1997. "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics*, 112(2): 443–477.
- Laibson, David. 2015. "Why Don't Present-Biased Agents Make Commitments." *American*

- Economic Review, Papers and Proceedings*, 105(5): 267–272.
- Laibson, David, Peter Maxted, Andrea Repetto, and Jeremy Tobacman. 2017. “Estimating discount functions with consumption choices over the lifecycle.” *Working Paper*.
- Machina, Mark J., and David Schmeidler. 1992. “A more robust definition of subjective probability.” *Econometrica*, 60(4): 745–780.
- Magnac, Thierry, and David Thesmar. 2002. “Identifying Dynamic Discrete Decision Processes.” *Econometrica*, 70(2): 801–816.
- Mahajan, Aprajit, and Alessandro Tarozzi. 2011. “Time Inconsistency, Expectations and Technology Adoption: The Case of Insecticide Treated Nets.” *Working Paper*.
- Mangasarian, Olvi L. 1994. *Nonlinear Programming. Classics in Applied Mathematics*, Society for Industrial and Applied Mathematics.
- Mas, Alexandre, and Enrico Moretti. 2009. “Peers at Work.” *The American Economic Review*, 99(1): 112–145.
- McClure, Samuel, David Laibson, George Loewenstein, and Jonathan Cohen. 2004. “Separate neural systems value immediate and delayed monetary rewards.” *Science*, 306: 503–507.
- Meier, Stephan, and Charles Sprenger. 2010. “Present-Biased Preferences and Credit Card Borrowing.” *American Economic Journal - Applied Economics*, 2(1): 193–210.
- Meier, Stephan, and Charles Sprenger. 2015. “Temporal Stability of Time Preferences.” *Review of Economics and Statistics*, 97(2): 273–286.
- O’Donoghue, Ted, and Matthew Rabin. 1999. “Doing it Now or Later.” *American Economic Review*, 89(1): 103–124.
- O’Donoghue, Ted, and Matthew Rabin. 2001. “Choice and Procrastination.” *The Quarterly Journal of Economics*, 116(1): 121–160.
- Polisson, Matthew, , John Quah, and Ludovic Renou. 2015. “Revealed preference tests over risk and uncertainty.” working paper.
- Rees-Jones, Alex. 2013. “Loss Aversion Motivates Tax Sheltering: Evidence from US Tax Returns.” *Working Paper*.
- Richter, Marcel K. 1966. “Revealed preference theory.” *Econometrica*, 34(3): 635–645.
- Rockafellar, Tyrrell. 1997. *Convex Analysis*. Princeton University Press.

- Rothschild, Michael, and Joseph E. Stiglitz. 1970. "Increasing risk: I. A definition." *Journal of Economic theory*, 2(3): 225–243.
- Rust, John. 1994. "Structural Estimation of Markov Decision Processes." In *Handbook of Econometrics Vol. 4.*, ed. Robert Engle and Daniel McFadden, 3081–3143. Amsterdam: New Holland.
- Sacerdote, Bruce I. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, 116(2): 681–704.
- Samuelson, Paul A. 1937. "A Note on Measurement of Utility." *The Review of Economic Studies*, 4(2): 155–161.
- Shang, Jen, and Rachel Croson. 2009. "A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods." *The Economic Journal*, 119(540): 1422–1439.
- Shapiro, Jesse. 2005. "Is there a daily discount rate? Evidence from the food stamp nutrition cycle." *Journal of Public Economics*, 89(2-3): 303–325.
- Shue, Kelly. 2013. "Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers." *The Review of Financial Studies*, , (6): 1401–1442.
- Slemrod, Joel, Charles Christian, Rebecca London, and Jonathan A. Parker. 1997. "April 15 Syndrome." *Economic Inquiry*, XXXV: 695–709.
- Souleles, Nicholas S. 1999. "The Response of Household Consumption to Income Tax Refunds." *American Economic Review*, 89(4): 947–958.
- Sprenger, Charles. 2015. "Judging Experimental Evidence on Dynamic Inconsistency." *American Economic Review, Papers and Proceedings*, 105(5): 280–285.
- Tanaka, Tomomi, Colin Camerer, and Quang Nguyen. 2010. "Risk and time preferences: Experimental and household data from Vietnam." *American Economic Review*, 100(1): 557–571.
- Varian, Hal R. 1983. "Nonparametric tests of models of investor behavior." *Journal of Financial and Quantitative Analysis*, 18(3): 269–278.
- von Neumann, John. 1953. "A certain zero-sum two-person game equivalent to the optimal assignment problem." *Contributions to the Theory of Games*, 2: 5–12.
- Zimmerman, David J. 2003. "Peer effects in academic outcomes: Evidence from a natural experiment." *Review of Economics and Statistics*, 85(1): 9–23.