



# Essays in Behavioral and Experimental Economics

## Citation

Taubinsky, Dmitry. 2014. Essays in Behavioral and Experimental Economics. Doctoral dissertation, Harvard University.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:12269825>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# Essays in Behavioral and Experimental Economics

A dissertation presented

by

Dmitry Taubinsky

to

The Department of Business Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

April 2014

© 2014 Dmitry Taubinsky

All rights reserved.

*Dissertation Advisor:*  
**David Laibson**

*Author:*  
**Dmitry Taubinsky**

## **Essays in Behavioral and Experimental Economics**

### **Abstract**

This dissertation consists of three essays examining the implications of human psychology for economic behavior and market outcomes.

The first chapter formalizes a model of people's inattention to choices and actions in dynamic decision environments. The model provides a new foundation for habit-forming behaviors, and also generates surprising implications for how deadline length can affect task completion rates. I report results from two "extra-lab" experiments that jointly vary reminder cues and payoff parameters to test and confirm the predictions of the model. I also embed this model of consumer behavior in various market settings and study how firms structure reminder advertising campaigns, consumer rebates, and free-trial offers.

The second chapter, coauthored with Hunt Allcott, examines the frequently made hypothesis that consumers are imperfectly informed about or inattentive to the energy costs associated with energy-using appliances. We study two field experiments that provide information on energy costs and product lifetimes for compact fluorescent lightbulbs vs. traditional incandescent bulbs. We then propose a general model of consumer bias in choices between energy-using durables, derive formulas for quantifying the welfare implications of such bias, and evaluate the implications of existing policies. Results suggest that moderate CFL subsidies may be optimal, but that imperfect information and inattention do not appear to justify a ban on traditional incandescent lightbulbs in the absence of other inefficiencies.

The third chapter, coauthored with Holger Herz, has its starting point in the now well-established fact that people's desire for fair transactions can play an important role in negotiations, organizations, and markets. In this chapter, we show that markets can also

shape what people consider to be a fair transaction. We propose a simple and generally-applicable model of *path-dependent fairness preferences*, in which past experiences shape preferences, and we experimentally test the model's predictions. We find that previous exposure to competitive pressure substantially and persistently reduces subjects' fairness concerns, making them more likely to accept transactions in which they receive a low share of the surplus. But consistent with our theory, we also find that past experience has little effect on subjects' inclinations to treat others unfairly.

# Contents

Abstract . . . . .	iii
Acknowledgments . . . . .	xiii
<b>1 From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Model and Evidence . . . . .	9
1.2.1 The Decision Environment . . . . .	9
1.2.2 Attention Dynamics . . . . .	10
1.2.3 Strategies and Beliefs . . . . .	12
1.2.4 Psychological foundations: Evidence and examples . . . . .	13
1.2.5 Remarks . . . . .	17
1.2.6 Relation to economics work on limited attention . . . . .	17
1.3 Repeated actions . . . . .	19
1.3.1 Theory . . . . .	19
1.3.2 Experimental Evidence . . . . .	26
1.4 Tasks With Deadlines . . . . .	34
1.4.1 Theory . . . . .	34
1.4.2 Experimental Evidence . . . . .	44
1.5 Inattention in the Market . . . . .	50
1.5.1 Rebates and Related Applications . . . . .	50
1.5.2 Optimal Cue Provision by an Interested Party . . . . .	57
1.6 Concluding Remarks . . . . .	63
1.6.1 Recap . . . . .	63
1.6.2 Extensions . . . . .	64
<b>2 The Lightbulb Paradox: Evidence from Randomized Experiments</b>	<b>67</b>
2.1 Introduction . . . . .	67
2.2 Background . . . . .	74
2.2.1 "The Lightbulb Paradox" . . . . .	74
2.2.2 Economic Reasons for Standards and Subsidies . . . . .	75

2.3	TESS Experiment . . . . .	77
2.3.1	Survey Platform . . . . .	77
2.3.2	Experimental Design . . . . .	78
2.3.3	Data . . . . .	84
2.3.4	Empirical Strategy and Results . . . . .	86
2.4	In-Store Experiment . . . . .	94
2.4.1	Experimental Design . . . . .	94
2.4.2	Data . . . . .	96
2.4.3	Empirical Strategy and Results . . . . .	98
2.5	A Framework for Policy Analysis . . . . .	100
2.5.1	Consumers . . . . .	100
2.5.2	The Policymaker . . . . .	101
2.5.3	First-Order Approximation to Optimal Subsidy . . . . .	107
2.6	Policy Evaluation . . . . .	108
2.6.1	Inferring Bias from Treatment Effects . . . . .	108
2.6.2	"Structural" Models of Bias . . . . .	109
2.6.3	Using the TESS Experiment Results . . . . .	111
2.6.4	Using the In-Store Experiment Results . . . . .	116
2.7	Conclusion . . . . .	117
<b>3</b>	<b>Market Experience is a Reference Point in Judgments of Fairness</b>	<b>120</b>
3.1	Introduction . . . . .	120
3.2	Experimental Design . . . . .	126
3.2.1	Phase 1: Market Games . . . . .	126
3.2.2	Phase 2: Ultimatum Game . . . . .	127
3.2.3	Procedures . . . . .	128
3.3	Theory and Hypothesis Development . . . . .	130
3.3.1	Set-up . . . . .	131
3.3.2	Equilibrium in the PC Market in Phase 1 . . . . .	132
3.3.3	Equilibrium in the RC Market in Phase 1 . . . . .	133
3.3.4	Phase 2 Behavior . . . . .	134
3.3.5	Convergence . . . . .	136
3.3.6	Discussion of Assumptions . . . . .	138
3.3.7	Testable Hypotheses . . . . .	138
3.4	Results . . . . .	140
3.4.1	Phase 1: The Effect of Competition on Offers and Acceptance Decisions	140
3.4.2	Phase 2: The Effect of Experience on Responder Behavior . . . . .	141
3.4.3	The Effect of Experience on Proposer Behavior . . . . .	144

3.4.4	Convergence of Fairness Preferences . . . . .	147
3.4.5	Discussion of Results . . . . .	148
3.5	Concluding Remarks . . . . .	151
<b>References</b>		<b>154</b>
<b>Appendix A Appendix to Chapter 1</b>		<b>170</b>
A.1	Extensions and Additional Examples . . . . .	170
A.1.1	Infinite Horizon . . . . .	170
A.1.2	Endogenous Cue Generation . . . . .	171
A.1.3	Modeling Partial Naivete . . . . .	174
A.1.4	Response elasticities for section 1.3.1 . . . . .	175
A.1.5	Microfoundations for the price floor assumption . . . . .	175
A.2	Daily Action Experiment: Supplementary Material . . . . .	176
A.2.1	Theoretical Extensions . . . . .	176
A.2.2	Additional Experimental Details . . . . .	177
A.2.3	Robustness to Demographic Controls . . . . .	178
A.2.4	Robustness to calendar date random effects . . . . .	179
A.2.5	Analysis of Post-Experimental Survey Results . . . . .	179
A.3	Task Completion Experiment: Supplementary Material . . . . .	181
A.3.1	Reminders Text . . . . .	181
A.3.2	Demographics . . . . .	182
A.3.3	Day of week effects . . . . .	183
A.3.4	Robustness . . . . .	184
A.3.5	Reminder Types . . . . .	185
A.4	Proofs of Mathematical Results . . . . .	185
A.4.1	Proofs for repeated action results . . . . .	185
A.4.2	Proofs for one time actions results . . . . .	193
A.4.3	Proofs for rebate market results . . . . .	203
A.4.4	Proofs for reminder advertising results . . . . .	205
<b>Appendix B Appendix to Chapter 2</b>		<b>209</b>
B.1	Details of TESS Experiment . . . . .	209
B.2	Additional TESS Results . . . . .	219
B.3	iPad Total Cost Comparison Screen . . . . .	226
B.4	Appendix to Theoretical Framework . . . . .	227



<b>Appendix C Appendix to Chapter 3</b>	<b>229</b>
C.1 Additional tables . . . . .	229
C.2 Proposer optimization . . . . .	230
C.3 Proofs of Propositions (online publication only) . . . . .	232

## List of Tables

1.1	Probability of Completing a Survey on any Given Day in Week 3 . . . . .	32
1.2	Fraction of Subjects Completing Task, by Experimental Condition . . . . .	47
1.3	Probability of Completing Task, by Experimental Condition . . . . .	47
2.1	Descriptive Statistics and Balance for TESS Experiment . . . . .	85
2.2	Effects of TESS Informational Interventions . . . . .	89
2.3	Perceived Intent of TESS Study . . . . .	91
2.4	Effects on Beliefs . . . . .	93
2.5	Effects on Important Factors in Purchase Decision . . . . .	94
2.6	Descriptive Statistics and Balance for In-Store Experiment . . . . .	97
2.7	Effects of In-Store Informational Intervention . . . . .	99
2.8	Welfare Analysis Using TESS Results . . . . .	114
3.1	Overview of Matching Groups . . . . .	129
3.2	Regression analysis . . . . .	144
3.3	Proposer Offers . . . . .	146
3.4	Proposers' beliefs about acceptance . . . . .	147
3.5	Time trends in minimum acceptance thresholds . . . . .	148
A.1	Demographics by Experimental Condition . . . . .	177
A.2	Replication of Table 1.1 with Demographic Controls. . . . .	178
A.3	Replication of Table 1.1 With Clustering at the Calendar Date Level . . . . .	179
A.4	Effect of Using Own Reminder Technology, by Experimental Condition. . . . .	180
A.5	Demographics by Experimental Condition . . . . .	182
A.6	Day of Week Effects . . . . .	183
A.7	Robustness to Day of Week and Demographics . . . . .	184
A.8	Different Types of Reminders Don't Have Differential Effects . . . . .	185
B.1	Association Between Individual Characteristics and CFL Demand . . . . .	221
B.2	Correlation of Treatment Effects with Self-Monitoring Scale . . . . .	222

B.3	Sensitivity of TESS Welfare Analysis to Assumed Mean Censored Values Assuming Top-Coded and Bottom-Coded WTPs Average \$12 and -\$12, Re- spectively . . . . .	223
B.4	Sensitivity of TESS Welfare Analysis to Assumed Mean Censored Values Assuming Top-Coded and Bottom-Coded WTPs Average \$20 and -\$20, Re- spectively . . . . .	224
C.1	P-Values of Wilcoxon Rank-Sum tests . . . . .	229
C.2	Minimum Acceptance Thresholds and Cognitive Reflection Test Scores . . .	230
C.3	Profit maximizing offers by matching groups . . . . .	230

## List of Figures

1.1	Timeline for Repeated Action Experiment . . . . .	29
1.2	Surveys Completed Per Week, by Experimental Condition . . . . .	31
1.3	“U-shaped” Curve . . . . .	41
1.4	Fraction of Subjects Completing Task on Each Day, by Condition . . . . .	49
2.1	CFL Demand Curves . . . . .	87
2.2	Histogram of Changes in WTP . . . . .	88
2.3	Marginal Internality Example . . . . .	104
2.4	Demand Curves . . . . .	107
2.5	TESS Experiment Treatment Effects by Initial WTP . . . . .	112
2.6	Welfare Calculation Using TESS Experiment . . . . .	113
2.7	Policy Analysis Using In-Store Experiment . . . . .	117
3.1	Average Offers and Acceptance Rates Over Time, Respectively. . . . .	141
3.2	Minimal Acceptance Thresholds of Responders. . . . .	142
3.3	Mean and Median Minimal Acceptance Thresholds of Responders, by Match- ing Group. . . . .	143
3.4	Mean and Median Offers of Proposers, by Matching Group. . . . .	145
B.1	Introductory Screen . . . . .	209
B.2	Baseline Lightbulb Choices (Top of Screen) . . . . .	210
B.3	Detailed Product Information . . . . .	210
B.4	Baseline Product Information (Bottom of Screen) . . . . .	211
B.5	Total Cost Information Screen . . . . .	211
B.6	Disposal and Warm-Up Information Screen . . . . .	212
B.7	Disposal and Warm-Up Information Screen . . . . .	212
B.8	Control Introductory Screen . . . . .	213
B.9	Number of Bulbs by Sector Information Screen . . . . .	213
B.10	Sales Trends Information Screen . . . . .	214
B.11	Endline Lightbulb Choices (Top of Screen) . . . . .	214
B.12	Endline Lightbulb Choices (Bottom of Screen) . . . . .	215

B.13 Post-Only Treatment Demand Curve . . . . .	225
B.14 iPad Total Cost Comparison Screen . . . . .	226

## Acknowledgments

Many people have contributed to this dissertation and the work that I've done over the past five years.

First, I would like to thank my advisers: David Laibson, Sendhil Mullainathan, Al Roth, and Mike Luca. David first introduced me to economic research nine years ago when I was only a freshman in college. Throughout the years, David has been a brilliant and caring mentor, fueling my journey with his enthusiasm and wisdom. My conversations with Sendhil, and his gift for flawlessly combining the intuitive and the abstract, have made me a more insightful researcher. I'm also lucky to have had many conversations with Al. His down-to-earth brilliance, wisdom, and unfaltering commitment to students have been an inspiration to me, both as an academic and as a human being. I'm thankful to Mike for his constant encouragement to connect economic research to simple real-world situations, as well as his infectious optimism.

For Chapter 1, I thank Hunt Allcott, John Beshears, Pedro Bordalo, Tom Cunningham, Drew Fudenberg, Ed Glaeser, Judd Kessler, Antonio Rangel, Alex Rees-Jones, Todd Rogers, Josh Schwartzstein, Andrei Shleifer, Tomasz Strzalecki, and Eric Zwick for helpful comments. I also thank the staff at the Harvard Decision Science Lab, especially Gabe Mansur, for technical assistance with the experiments reported in this paper. Finally, I thank the Harvard Business School Doctoral Office for financial support of this project.

For Chapter 2, I am first and foremost grateful to my coauthor Hunt Allcott. We are also grateful to Lucas Davis, Stefano DellaVigna, Mushfiq Mubarak, Sendhil Mullainathan, Emmanuel Saez, and other colleagues, as well as seminar audiences at Berkeley, Harvard, the NBER Public Economics Meetings, Stanford, the Stanford Institute for Theoretical Economics, and the University of California Energy Institute for constructive feedback. We thank our research assistants - Jeremiah Hair, Nina Yang, and Tiffany Yee - as well as management at the partner company, for their work on the in-store experiment. Thanks to Stefan Subias, Benjamin DiPaola, and Poom Nukulki at GfK for their work on the TESS experiment. We are grateful to the National Science Foundation and the Sloan Foundation

for financial support.

For Chapter 3, I am first and foremost grateful to my coauthor Holger Herz. We also thank Philippe Aghion, Ernst Fehr, Oliver Hart, Michael Kosfeld, David Laibson, Michael Luca, Sendhil Mullainathan, Al Roth, Klaus Schmidt, seminar participants at Frankfurt, Harvard and Zurich, as well as participants at the IMEBE meeting 2013 in Madrid, at the THEEM meeting 2013 in Kreuzlingen and at the ESA World meeting in Zurich 2013 for helpful comments and discussions. We thank the Harvard Law School Program on Negotiation, the Harvard Business School doctoral program, and the Center for Foundations of Economic Preferences at the University of Zurich for financial support of the project.

Finally, I would like to thank my family for their constant support, for encouraging me to pursue my passions, and for always being ready to impart wisdom and perspective. I thank my parents, Leonid and Svetlana, for always being there for me in tough times and happy times; my grandparents, Rimma, Adolph, and Dora, for sharing a lifetime of wisdom with me; my sister, Nika, for always being my best friend; and my soon to be wife and always partner in crime, Lauren, for always challenging me to be myself and to be passionate about what I do. I dedicate this dissertation to my family.

To my family



## Chapter 1

# From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice

### 1.1 Introduction

The relationship between people's preferences and actions is characterized by a variety of frictions. Frictionless theories of intertemporal choice are hard to reconcile with people's poor adherence to medication regimens, lack of exercise, low savings rates, frequently late bill payments, and inefficient energy use. Incomplete information, systematically biased beliefs, and self-control problems are often invoked to explain what appear to be suboptimal choices, and have been formally modeled in a variety of economic settings.

This paper theoretically and experimentally investigates a different bias that may play an important role in intertemporal decision-making: people are *inattentive*, and will not take an action that is not on their minds. At one point in time, a person may have fully intended to take all of his daily medications, pay his bills on time, and apply for a rebate; at another point in time, those once salient intentions may no longer be contemplated, buried by daily distractions, worries, and new goals.

In the case of poor compliance with medication regimens, for example, chronic disease patients frequently report “forgetting” as the reason for why they do not take all of their prescribed daily medication (MacDonell et al., 2013; Vyankandondera et al., 2013). This is corroborated by evidence from cognitive psychology experiments, which show that attentional and memory failures measured in controlled laboratory environments are significantly correlated with poor medication adherence (Poquette et al., 2013; Zogg et al., 2012).

Another growing body of evidence builds on a basic implication of limited attention: cues directing people’s attention to a particular action should make its execution more likely. In decisions affecting health outcomes, simple cues such as email or text message reminders have been shown to increase compliance with daily medication regimens (Vervloet et al., 2012), scheduling of check-ups and screening tests (DeFrank et al., 2009), exercise (Calzolari and Nardotto, 2012), and healthy food choices (Patrick et al., 2009). In financial decision making, simple cues have been shown to increase savings (Karlan et al., 2012a) and timely loan repayments (Cadena and Schoar, 2011). And various cues directing attention to electricity use have decreased energy consumption by increasing simple behaviors such as turning off energy-using appliances (Gilbert and Zivin, 2013; Jessoe and Rapson, forthcoming).

However, the full implications of inattention to choices and actions in dynamic decisions are poorly understood:<sup>1</sup> What makes a certain choice or action more or less likely to come to mind? When should reminder cues should be more or less effective, and how do they interact with financial incentives?<sup>2</sup> How do sophisticated, profit-maximizing firms respond to consumer inattention?

This paper proposes a dynamic model of *inattentive choice* and theoretically and exper-

---

<sup>1</sup>In contrast, the implications of limited capacity to update expectations in dynamic environments have been much more studied in rational inattention models and other approaches.

<sup>2</sup>Note that models of rational inattention (e.g. Sims 2003), which I discuss in more detail in Section 1.2.6, do not naturally allow for simple cues such as reminders to affect choice, and thus do not shed light on these issues. The simple reason is that an uninformative reminder should not lower the costs associated with processing information about a particular option.

imentally investigates its implications. The first ingredient of the model is that cues can affect behavior by reminding the person about the choice or action. The second ingredient is what research in psychology and other cognitive sciences calls *rehearsal* or *accessibility bias*: the idea that recent engagement with a certain choice or action will increase the likelihood that it rises to the top of the mind again. Two forms of rehearsal are key to the decisions studied in this paper: a weak form of rehearsal is when the person thinks about the action, a stronger form is when the person performs the action.

For repeated actions—such as taking medication, attending the gym, or turning off electricity-using appliances to cut energy use—rehearsal implies that these behaviors will be habit-forming. The more a person has performed these behaviors in the past, the more likely they are to be top of mind, and thus the more likely they are to be performed again. A consequence of rehearsal, therefore, is that temporary incentives have a ripple effect: for example, incentivizing people to make energy conserving actions a more integral part of their routines will increase the attention devoted to those actions in a way that will persist when those incentives are removed. I show that rehearsal also has implications for tasks that must be completed by a deadline; in particular, it can significantly diminish the option value of longer deadlines. A consumer who purchases a product with the intention of mailing in the accompanying \$50 rebate will be likely to think about the rebate again several days later but, because of rehearsal, will be less likely to think about it again in 4 weeks. As a consequence, consumers who are naive about their future inattention can be hurt by longer deadlines.

These predictions are consistent with a broad set of empirical findings, and show that patterns of behavior that are sometimes attributed to other theories such as hyperbolic discounting (Laibson, 1997; O'Donoghue and Rabin, 1999) or habit-forming *preferences* (Becker and Murphy, 1988) may also be a consequence of time-varying attention. At the same time, the model is sharply distinguished from other theories by its predictions about how changes in cues can modify behavioral responses to changes in payoff parameters. For tasks with deadlines, the model predicts that a small set of reminders can alter the

potentially perverse effect of longer deadlines. For repeated actions, the model makes predictions about how reminders can amplify or diminish the spillover effects of temporary shocks to peoples' routines. These novel predictions about the interaction of cues and payoff parameters are immediately translatable into new experimental designs, and are tested and confirmed in two experiments reported in this paper.

These results, as well as others, lead to four main contributions. First, I show that a simple and psychologically grounded model of inattentive choice provides a unifying explanation of what appear to be "non-standard" responses to changes in payoff parameters. Second, I show that the model extends the economic analysis of incentives by providing a framework for studying the effects of cue provision with a theoretical precision that generates sharp, testable hypotheses. The model generates novel predictions for the previously unexamined question of how cues and incentives interact to influence behavior, and provides a needed organization of conditions under which cues should or should not have a significant impact. Third, I test the new predictions of the model in two real-effort experiments that employ  $2 \times 2$  factorial designs to vary both payoff parameters and cues. These experiments provide the first direct evidence that both the deadline and the habit effects predicted by the rehearsal property of the model are, indeed, a consequence of time-varying attention. Fourth, I show that beyond its predictions for individual behavior, the inattention model can be fruitfully incorporated into a variety of market analyses. I illustrate the model's applicability by using it to theoretically study marketing tactics such as consumer rebates, and to complement existing theories of advertising by deriving optimal policies of *reminder advertising*.

Section 1.2 presents the formal model, which builds on the imperfect recall model in Mullainathan (2002), and related approaches by Ericson (2010) and Holman and Zaidi (2010). I consider a decision maker (DM) who each period must choose whether or not to take a certain action when it is available. But with some probability, the DM is *inattentive*, and therefore does not even consider taking the action. Closely following evidence from psychology, I assume that cues and rehearsal increase the probability of the DM being attentive. To complete the model and apply it to dynamic decisions, I further draw on the

experimental literature to specify how people form beliefs about the possibility of their future inattentiveness. Because the evidence discussed in Section 1.2 suggests that, on average, people vastly underestimate the likelihood of future inattentiveness, I draw out the additional implications of this naivete by considering two different ways of completing the model: *sophisticated beliefs* that are correct about future (in)attentiveness or *naive beliefs* that assume full attentiveness in all future periods.<sup>3</sup>

Section 1.3 studies the behavior of an inattentive DM in a repeated action setting such as medication adherence, exercise, or residential energy use. Section 1.3.1 contains the theoretical results, while Section 1.3.2 reports an experiment testing the new predictions. Section 1.3.1 shows that because engaging in a behavior increases the probability that it will be attentively considered in the future, the model predicts behavioral patterns similar to those predicted by preference-based theories of habit formation (Becker and Murphy, 1988), and documented for exercise (Charness and Gneezy, 2009; Acland and Levy, 2011) and residential energy use (Jesoe and Rapson, forthcoming). But at the same time—and consistent with evidence from gym attendance (Calzolari and Nardotto, 2012) and energy use (Allcott and Rogers, 2012; Gilbert and Zivin, 2013)<sup>4</sup>—the model predicts that cues such as reminders affect both current *and* future behavior.<sup>5</sup> Moreover, in addition to capturing both the habit effects and the cue effects in a single parsimonious framework, the model makes predictions about when increasing cues will amplify or diminish the impacts of incentives. Consistent with Jesoe and Rapson (forthcoming), the model predicts that people will be more responsive to incentives following a prior or concurrent increase in cues. At the same time, the model also predicts that the effects of current incentives on *future* behavior are decreasing in the strength of future cues. This second comparative static follows from

---

<sup>3</sup>Appendix A.1.3 further generalizes these two illustrative extremes and proposes a model of partial naivete.

<sup>4</sup>The cues described in Allcott and Rogers (2012), however, may act through channels other than just attention, because they also included normative information.

<sup>5</sup>This prediction resembles models of cue-triggered cravings (Laibson, 2001; Bernheim and Rangel, 2004). The scopes are very different, however. Laibson (2001) and Bernheim and Rangel (2004) focus on conditioned, visceral responses to immediate consumption opportunities, and apply the models mostly to addictions. Most of the repeated-action behaviors studied in this paper, in contrast, are “good” habits that typically involve temporally distant benefits that should not elicit any sort of craving.

core idea of the model that people do not need reminders for behaviors that are routine. That is, the model predicts that habits and reminders are substitutes.

The experiment reported in Section 1.3.2 examines the interaction between repeat performance and reminders. The task in this experiment resembles taking daily medication and involves completing a daily survey in return for \$1 per completed survey. The  $2 \times 2$  design varies 1) whether this task is available each day for three weeks straight or whether it is available for only the first and third weeks, and 2) whether or not participants receive daily reminders in the third week. I find that even without week 3 reminders, participants who are assigned to perform the task for three weeks straight complete approximately 5 out of 7 surveys in week 3. In contrast, those participants who experience the week 2 interruption and don't receive week 3 reminders complete only about 2.5 surveys. But consistent with the prediction that habits and reminders are substitutes, this effect of the week 2 interruption is almost entirely eliminated when subjects receive reminders in week 3. Reminders have only a minor effect on subjects who do not experience the interruption, but nearly double the week 3 completion rate of subjects who do experience the interruption. This interaction between repeat performance and reminders provides strong evidence that repeat performance of a behavior increases the likelihood that it is attentively considered in the future.

Section 1.4 studies the behavior of an inattentive DM who must complete some task—e.g., pay a bill, schedule a medical appointment, redeem a rebate, return a product—by a certain deadline. Section 1.4.1 contains the theoretical results, while Section 1.4.2 reports an experiment testing the new predictions. The central theoretical results in Section 1.4.1 concern how extending the deadline affects completion rates and welfare. On the one hand, the model predicts that a longer deadline increases a *sophisticated* DM's welfare. On the other hand, the model predicts that a longer deadline can decrease a *naive* DM's welfare and completion probability *when the likelihood of being attentive to the task decays over time*. Intuitively, this is because the naive DM does not realize that he will be significantly less likely to think of the task at a later point in time, and thus puts off the task more than

he should with a longer deadline. This result is consistent with experimental evidence on short vs. long deadlines for rebate redemption (Silk, 2004), gift certificate redemption (Shu and Gneezy, 2010), and product returns (Janakiraman and Ordóñez, 2012). The second central theoretical prediction for task completion helps differentiate the inattention mechanism from alternative explanations of deadline effects. The prediction is that *if* a longer deadline leads to lower completion rates *then* an appropriately placed reminder will have a significantly larger effect on decision makers facing the longer deadline, potentially reversing the “perverse” effect of the longer deadline.

Section 1.4.2 reports an experiment testing the new prediction about the interaction of deadlines and reminders. Participants in the experiment receive a cash reward for completing a 20 minute questionnaire by a certain deadline. The  $2 \times 2$  design varies the deadline length between either 2 days or 21 days, and varies whether participants receive reminders during the last two days including the deadline. Replicating previous evidence, I find that among participants who do not receive reminders, the longer deadline *decreases* the probability of completing the task from 59% to 42%. However, I find that reminders increase completion rates by 15 percentage points for subjects facing the shorter deadline, and increase completion rates by a striking 31 percentage points for subjects facing the longer deadline. Thus, while existing results on deadline effects replicate in the absence of reminders, they are nearly eliminated by reminders—a prediction unique to the inattention mechanism.

Section 1.5 builds on the theoretical results in Sections 1.3 and 1.4 to formally explore some market implications of the model. Section 1.5.1 builds on Section 1.4 and proposes a formal model of consumers rebates. In contrast to previous theoretical work—which has focused on price discrimination and has modeled consumers’ redemption decision as a static, one-period choice—I consider a dynamic model of the rebate redemption process and use the inattention model to analyze the previously ignored question of how firms choose the redemption deadlines. Different from existing theoretical work, but consistent with policymakers’ and industry experts’ claims, the model predicts that rebates are deceptively

attractive to inattentive and naive consumers. Moreover, the model predicts that firms will use “intermediate length” deadlines to maximally exploit consumer mistakes. Finally, the model shows that consumer rebates can facilitate socially inefficient transactions, and provides a formal framework for evaluating various policy proposals to regulate rebate offers. Building on the insights developed in the rebate model, I also discuss implications for product return policies and automatic renewal billing.

Section 1.5.2 builds on Section 1.3 to study the optimal cue provision strategy of an organization interested in increasing consumers’ likelihood of taking some repeated action. The simple advertising model developed in this section generates insights for a variety of different types of communications: A health care provider or insurer using SMS messages or phone calls to remind chronic disease patients to take their medications, an organization sending reports to remind consumers of ways to save energy, or a firm advertising a repeat-purchase product to make sure that it stays top of mind. I show that as in models of informative advertising, the returns to each additional message are decreasing in the intensity of previous communications. But in contrast to models of informative advertising, the inattention model predicts that the optimal intensity of reminder messages should not converge to zero in the long run, and that the optimal advertising strategy might involve cycles or intermittent messages. Moreover, when the behavioral rehearsal effect is sufficiently strong, the model also predicts that the optimal message intensity will be non-monotonic in consumers’ preferences for choosing the action; that is, reminder messages will be most effective for behaviors that consumers want to take, but not so often that these actions become habitual. These results, as well as others, illustrate how the model complements existing theories of informative and persuasive advertising by providing new insights and foundations for what marketers refer to as reminder advertising.

Section 1.6 discusses some limitations of the analysis in this paper, as well as directions for future research.



## 1.2 Model and Evidence

Sections 1.2.1-1.2.3 set up the model, Section 1.2.4 further motivates the assumptions, Section 1.2.5 offers some further remarks on the model, and Section 1.2.6 briefly reviews the related economics literature.

### 1.2.1 The Decision Environment

I define the model for a general binary choice decision environment in which payoffs and the action space can depend arbitrarily on the past history of actions. This encompasses actions that must be taken only once over the course of  $T$  periods, such as paying a bill, and actions that are taken repeatedly every period, such as taking medication. Although not analyzed in this paper, the model can also be applied to actions that are taken “every so often”; e.g., actions that must be taken once every five periods.

There are  $T$  periods  $t = 0, \dots, T$ , with  $T < \infty$ . Each period the decision maker (DM) makes a choice  $x_t$  from a choice-set  $X_t(h_t) \in \{\{d\}, \{d, a\}\}$ , where  $h_t = (x_0, \dots, x_{t-1})$  denotes the history of choices up to period  $t$ . The choice  $x_t = a$  is an *active* choice—e.g., paying the bill or taking the medication. The choice  $x = d$  is a passive default—e.g., not paying the bill or not taking the medication. When  $X_t(h_t) = \{d\}$ , there is no choice to be made in period  $t$ . In the bill example, for instance, this will be the case if the bill has already been paid in period  $\tau < t$ .

The DM’s period  $t$  flow utility after choice  $x_t$  is given by  $u_t(x_t, \zeta_t, h_t)$ , where the  $\zeta_t$  are independent random draws from a distribution  $F$  with bounded support in  $[\underline{\zeta}, \bar{\zeta}]$ , and are realized prior to the DM taking his action. Unless otherwise stated, I will assume that  $F$  is atomless and fully supported on  $[\underline{\zeta}, \bar{\zeta}]$ . In this case, I will assume that  $F$  has a density function, which I will denote by  $f$ .

Variation in  $\zeta_t$  represents fluctuations in daily opportunity costs or variation in taste shocks. I assume that  $u(d, \zeta_t, h_t) = 0$  for all  $\zeta_t$  and  $h_t$ , so that only the utility from the non-default choice fluctuates. The DM’s period  $t$  utility from a sequence of realizations  $u_t, u_{t+1}, \dots$  is given by  $U_t = \sum_{\tau \geq t} u_\tau$ . Note that for simplicity, I assume no discounting,

though allowing for an exponential discount factor would not change the results in any way. Appendix A.1.1, generalizes the model to infinite horizons and time discounting.

## 1.2.2 Attention Dynamics

Each period, the DM either thinks about  $a$ , denoted  $\alpha_t = 1$ , or does not think about  $a$ , denoted  $\alpha_t = 0$ . When  $\alpha_t = 0$ , he always chooses the default  $d$ . When  $\alpha_t = 1$ , the DM compares  $a$  to the default alternative and chooses the better option subject to the DM's beliefs about his own future actions. I will describe the state  $\alpha_t = 1$  as *attentive* and the state  $\alpha_t = 0$  as *inattentive*.

Each period, the DM also receives a set of attention cues  $\Omega_t \sim G_t$ , with strength  $\sigma(\Omega_t) \in [0, 1]$ . Cues increase the probability that the DM is attentive, and may include a host of events that direct attention to the choice of  $x_t = a$ , including advertisements, fliers, email or text message reminders, conversations with others, and various visual and auditory cues intended to serve as reminders. I will frequently use the shorthand of writing  $\sigma$  instead of  $\sigma(\Omega_t)$ , and I will let  $H_t$  denote the period  $t$  distribution of values of  $\sigma$ .

In period  $t \geq 1$ , the DM is more likely think about  $a$  if 1) he has thought about it in period  $t - 1$  (i.e.  $\alpha_t = 1$ ); if 2) he actually engages in the behavior  $a$  in period  $t - 1$  (i.e.,  $x_{t-1} = a$ ); and if 3) he receives a salient set of cues in period  $t$ . I will refer to the effect of  $\alpha_{t-1}$  on period  $t$  attentiveness as *mental rehearsal* and I will refer to the effect of  $x_{t-1}$  on period  $t$  attentiveness as *behavioral rehearsal*. Section 1.2.4 reviews the cognitive psychology literature that motivates these attention effects.

Formally, I assume that for  $t \geq 1$ ,  $Pr(\alpha_t = 1) = g(\alpha_{t-1}, x_{t-1}, \sigma(\Omega_t))$ , where  $(\alpha_t, x_t) \in \{(0, d), (1, d), (1, a)\}$  and  $g$  satisfies the following assumptions:<sup>6</sup>

**A1**  $g$  is continuous and increasing in  $\sigma$ , and  $g(0, d, 1) = 1$

**A2**  $g(1, d, 0) > 0$

**A3**  $g(0, d, \sigma) \leq g(1, d, \sigma) \leq g(1, a, \sigma)$  for all  $\sigma$

---

<sup>6</sup>Note that since the DM cannot choose  $x_t = a$  when he is inattentive, the pair  $(0, a)$  is not possible.

**A4** For any  $\sigma_1 < \sigma_2$ ,  $g(0, d, \sigma_2) - g(0, d, \sigma_1) \geq g(1, d, \sigma_2) - g(1, d, \sigma_1) \geq g(1, a, \sigma_2) - g(1, a, \sigma_1)$

The first assumption states that the stronger the set of cues the DM receives, the more likely he is to be attentive. Moreover, there exists a sufficiently strong set of cues that would make the DM fully attentive. The second assumption states that if the DM was attentive last period, there is always a chance that he will be attentive this period. The third assumption formalizes the idea that rehearsal increases the likelihood of being attentive to  $a$ . The fourth assumption states that the effect of cues on period  $t$  attentiveness is less pronounced if behavioral and mental rehearsal already make  $a$  more top of mind.

The fourth assumption can be motivated with the following intuitive model of cues: Suppose that the DM receives a set of cues  $\Omega = \{\omega^1, \dots, \omega^K\}$  in period  $t$ . The effect of each additional cue  $\omega_k$  is that it triggers thoughts of the action with probability  $\sigma(\{\omega_k\})$ , and thus decreases the probability of being inattentive by a factor  $1 - \sigma(\{\omega_k\})$ . Thus the set of cues  $\Omega$  makes the DM attentive with probability

$$g(\alpha_{t-1}, x_{t-1}, \Omega) = 1 - [(1 - g(\alpha_{t-1}, x_{t-1}, 0))(1 - \sigma(\{\omega_1\})) \cdots (1 - \sigma(\{\omega_K\}))] \quad (1.1)$$

and can be said to have an aggregate strength of  $1 - \prod_{i=1}^K (1 - \sigma(\{\omega_i\}))$ . Clearly, specification (1.1) satisfies A1-A4 as long as  $g(\alpha_{t-1}, x_{t-1}, 0)$  satisfies A3.

Assumptions A1-A4 will be in full force throughout the paper. Some results will also rely on strengthening A3 and A4 to require that rehearsal strictly matters:

**A3'**  $g(0, d, \sigma) < g(1, d, \sigma) < g(1, a, \sigma)$  for all  $\sigma \in [0, 1)$ .

**A4'** For any  $\sigma_1 < \sigma_2$ ,  $g(0, d, \sigma_2) - g(0, d, \sigma_1) > g(1, d, \sigma_2) - g(1, d, \sigma_1) > g(1, a, \sigma_2) - g(1, a, \sigma_1)$ .

Note that specification (1.1) satisfies A4' when  $g(\alpha_{t-1}, x_{t-1}, 0)$  satisfies A3'.

In period 0,  $Pr(\alpha_0 = 1) = g(0, d, \sigma(\omega_0))$ , with  $\sigma(\omega_0)$  potentially very close to 1 if that's when the DM first learns about the behavior option. To further simplify notation, I will set  $\gamma_t(\alpha_{t-1}, x_{t-1}) = \int g(\alpha_{t-1}, x_{t-1}, \sigma) dH_t(\sigma)$ , and set  $\gamma_0 = \int g(0, d, \sigma) dH_0(\sigma)$ . That is,  $\gamma_t(\alpha_{t-1}, x_{t-1})$  is simply the probability of being attentive in period  $t$ , as a function of period  $t - 1$  events.

### 1.2.3 Strategies and Beliefs

#### Sophisticated Decision Makers

A sophisticated DM correctly anticipates the dynamics of his (in)attention, and optimally chooses his actions with those dynamics in mind. His strategies can be computed by backward induction.

To formalize, let  $E_F$  denote the expectation taken with respect to  $F$ , and let  $(h_t, x_t)$  denote the period  $t + 1$  history that results when  $x_t$  is chosen after history  $h_t$ . Let  $V_t^s(h_t, \alpha_t)$  denote the sophisticated DM's expected period  $t$  utility conditional period  $t$  history  $h_t$  and on whether or not he is attentive in period  $t$ . Let  $\mathbf{x}^s(h_t, \zeta_t)$  denote the DM's period  $t$  strategy conditional on being attentive. The functions  $V_t^s$  and  $\mathbf{x}^s$  are defined recursively below. To ease notation, I sometimes suppress the arguments of  $\mathbf{x}^s(\cdot)$ . For  $t = T$ ,

$$V_T^s(h_T, \alpha_t) = \alpha_t E_F u(\mathbf{x}^s, \zeta_T, h_T) \quad (1.2)$$

$$\mathbf{x}^s(h_T, \zeta_T) = \operatorname{argmax}_{x \in A(h_t)} u(x, \zeta_T, h_T) \quad (1.3)$$

and for  $t < T$ ,

$$V_t^s(h_t, \alpha_t) = \alpha_t E_F [u(\mathbf{x}^s, \zeta_t, h_t) + V_{t+1}^s((h_t, \mathbf{x}^s), \gamma(1, \mathbf{x}^s))] + (1 - \alpha_t) E_F V_{t+1}^s((h_t, d), \gamma(0, d)) \quad (1.4)$$

$$\mathbf{x}^s(h_t, \zeta_t) = \operatorname{argmax}_{x \in A(h_t)} \{u(x, \zeta_t, h_t) + V_{t+1}^s((h_t, x), \gamma(1, x))\} \quad (1.5)$$

As formalized in equation (1.5), a sophisticated DM recognizes that he may be inattentive in the future and considers how cues and rehearsal affect the probability of being attentive in period  $t + 1$ .

#### Naive Decision Makers

As summarized in Section 1.2.4, many people appear to be naive about the possibility of future inattention. To investigate the role of naivete in a stark and illustrative way, in the body of the paper I make the extreme assumption that a naive DM is *fully naive*; that is,

he thinks that he will be perfectly attentive in the future.<sup>7</sup> In appendix A.1.3, however, I introduce models of partial naivete. In general, a sufficiently—though not fully—naive DM would exhibit many of the same forms of misoptimization that a fully naive DM exhibits, but to a lesser extent.<sup>8</sup>

Conditional on being attentive, the naive DM makes the same choice that a fully attentive DM would make. And the utility that a naive DM *thinks* he will realize, in expectation, corresponds to the utility that a perfectly attentive DM *actually* realizes, in expectation.

Formally, let  $\mathbf{x}_t^n$  denote the naive DM's strategy, and let  $\tilde{V}_t^n$  denote the naive DM's expectation of his time  $t$  utility conditional on history  $h_t$ . Then analogous to equations (1.2)-(1.5),

$$\tilde{V}_T^n(h_T) = E_F u(\mathbf{x}_T^n, \zeta_T, h_T) \quad (1.6)$$

$$\mathbf{x}_T^n(\zeta_T) = \operatorname{argmax}_{x \in A(h_T)} u(x, \zeta_T, h_T) \quad (1.7)$$

and for  $t < T$ ,

$$\tilde{V}_t^n(h_t) = E_{F_t} [u(\mathbf{x}_t^n, \zeta_t, h_t) + \tilde{V}_{t+1}^n((h_t, \mathbf{x}_t^n))] \quad (1.8)$$

$$\mathbf{x}_t^n(h_t, \zeta_t) = \operatorname{argmax}_{x \in A(h_t)} \{u(x, \zeta_t, h_t) + \tilde{V}_{t+1}^n((h_t, x))\} \quad (1.9)$$

The expected period  $t$  utility  $V_t^n$  that a naive DM actually realizes is determined analogously to equations (1.2) and (1.4), with  $\mathbf{x}^n$  in place of  $\mathbf{x}^s$  and with  $V_t^n$  in place of  $V_t^s$ .

#### 1.2.4 Psychological foundations: Evidence and examples

In this section I discuss psychological evidence motivating the foundations of this model. Readers may skip this section without loss of continuity.

---

<sup>7</sup>This is analogous to O'Donoghue and Rabin's (1999) analysis of (fully) naive versus sophisticated hyperbolic discounters.

<sup>8</sup>In the decisions studied in Section 1.4, for example, the partially naive DM would still put off the task more than he should, but not as much as the fully naive DM. And as I discuss in appendix A.1.3, the results of Section 1.3 for naive DMs go through verbatim under milder forms of naivete.

## Attention dynamics

*Rehearsal.* A key difference between one-time actions such as redeeming a rebate and repeated actions such as taking medication is that repetition of a behavior option can increase the likelihood that it will come to mind again. The idea that future retrieval of information or intentions is made easier by rehearsing the retrieval has a long tradition of research in cognitive psychology (Atkinson and Shiffrin, 1969). In laboratory studies, Jacoby et al. (2001) coined the term “accessibility bias” to refer to the effect that repetition of a response has on the ease with which it comes to mind.<sup>9</sup> Recent work has proposed that repeated engagement in recycling behaviors (Tobias, 2009) or repeated purchasing of a product (Henderson et al., 2011) increases the attention devoted to those actions.

Other work (Sellen et al., 1997) suggests that even passive thoughts about one’s intentions can increase the likelihood that they will be retrieved in the future. As Sellen et al. (1997) argue, however, mental rehearsal is not enough to permanently preserve intentions at the top of the mind in the absence of external cues. McBride et al. (in press), for example, gave study participants postcards and instructed participants to mail them back—in return for a \$50 lottery—after randomly chosen delays of 1, 2, 5, 14, or 30 days. Corroborating the hypothesis that mental rehearsal does not by itself preserve goals, performance was significantly better by subjects randomized into the short delays.<sup>10</sup> Analogous results on task completion have been obtained for laboratory tasks with delays varying between 5 minutes and 60 minutes (McBride et al., 2011; Martin et al., 2011).<sup>11</sup>

---

<sup>9</sup>In laboratory tests of accessibility bias (Jacoby et al., 2001; Hay and Jacoby, 1996), subjects first become accustomed to responding a certain way to stimulus words such as “knee.” For some subjects the required response to “knee” is “bone” 75% of the time while for others it is “bend.” The next day, subjects are asked to memorize a list of word pairs such as “knee/bone.” In the quiz stage of day 2, subjects who become accustomed to responding to “knee” with “bone” in day 1, are both much faster and more accurate in their responses to “knee” in day 2 than are subjects who become accustomed to the response “bend.”

<sup>10</sup>An average 65% return rate for delays of 1, 2, or 5 days versus an average 48% return rate for delays of 14 or 30 days.

<sup>11</sup>However, it is not clear that these laboratory tasks tap into the same psychological mechanisms as the naturalistic experiments. The typical “prospective memory” tasks require subjects to respond a certain way to a cue or prompt after a short time delay. Success in these tasks relies on short bursts of vigilance efforts that don’t seem to have direct “real world” analogs. Most field behaviors involve unprompted actions, such as “canceling the gym membership at some point next week.”

*Cues.* Many types of cues may trigger thoughts of the behavior. Visual cues such as noticing the pill bottle can trigger thoughts of the associated action; glancing at a billboard can refresh consideration of the advertised product. Auditory cues can have similar effects. Charles et al. (2007) found that a combination of visual and audio reminders built into an electronic pill bottle had a substantial impact on adherence to inhaled corticosteroid therapy over the course of 24 weeks.

Firms and organizations often capture attention through various messages: reminder emails, SMS messages, or advertisements. Recent evidence from randomized controlled trials shows that reminder messages have large effects on health behaviors such as gym attendance (Calzolari and Nardotto, 2012), sunscreen use (Armstrong et al., 2009), adherence to medication regimens (Krishna et al., 2009; Vervloet et al., 2012), obtaining immunizations (Szilagyi et al., 2000), scheduling/attending medical appointments (Altmann and Traxler, 2012), and even weight loss (Patrick et al., 2009).<sup>12</sup> In financial settings, Karlan et al. (2012a) find that reminders increase savings deposits, while Cadena and Schoar (2011) find that reminders increase timely loan repayments by small business owners.<sup>13</sup> In marketing, Nedungadi (1990) and Mitra and Lynch. (1995) find that non-persuasive product primes can increase the consideration and subsequent purchase of the product, without increasing the product's perceived value.

Finally, incidental events such as conversations with others can also constitute cues. Hearing a colleague complain about her referee report, for example, may trigger thoughts about one's own refereeing duties.

*Cues are imperfect.* Although cues can have powerful effects on behavior, most cues are still imperfect in the sense that a period  $t$  cue cannot guarantee attentiveness in period  $t$ . One reason for this imperfection is that emails, SMS messages, fliers, calendar reminders,

---

<sup>12</sup>Over the course of 4 months, Patrick et al. (2009) found that individuals receiving SMS messages reminding them to purchase healthy foods, pack healthy snacks, drink water, and other such activities lost 1.97kg more than the control group. Haapala et al. (2012) found that over the course of 12 months, individuals receiving SMS reminders lost 3.4kg more than the control group ( $N = 125$ ).

<sup>13</sup>See, however, Karlan et al. (2012b) who find that only personalized SMS reminder messages seem to affect repayments of loans in the Philippines.

and many other types of communications do not always reach the targeted person. There is no guarantee that the person will check his mail or phone, or be at his computer. The person may also choose to discard many of these communications before even reading them. Second, even within a time period of 1 day, a person will not necessarily receive the cue at a time when he is ready to act (Tobias, 2009). A reminder email to schedule a medical appointment might be read by the person when he is at work, but then subsequently forgotten several hours later when the person is ready to act. Third, indirect cues such as conversations with others are not guaranteed to trigger associated thoughts of the behavior. Fourth, a person with a limited capacity to process information may not meaningfully process all communications or stimuli. Section 1.6, for example, discusses how a person may become desensitized to certain cues.

### **Naivete**

Experimental studies show that people overestimate the likelihood that they will take a certain action on a later date. Ericson (2011), for example, used an incentive compatible mechanism to elicit subjects' beliefs about the likelihood that on a later date, they will send an email to the experimenter to claim a \$20 payment (that would subsequently be mailed to them). Quite strikingly, only 53% of the subjects sent the email, though the average incentive-compatible forecast was 76%. While the failure to carry out such a high reward/low cost task suggests significant inattention,<sup>14</sup> the large discrepancy between forecasts and actual behavior also suggests significant naivete. In similar tasks resembling rebate redemption, Silk (2004), Letzler and Tasoff (2013),<sup>15</sup> and Shu and Gneezy (2010) also find a large and robust discrepancy between forecasts and behavior.

---

<sup>14</sup>Ericson's (2011) calibrations show that alternative explanations such as time inconsistency or underestimation of effort cannot explain subjects' behavior given the high benefit-cost ratio.

<sup>15</sup>Letzler and Tasoff (2013) also found that a reminder sent approximately 2.5 weeks before the deadline reduced overconfidence by 7 percentage points by increasing redemption, though the difference was not statistically significant. Results in section 1.4, however, suggest that reminders sent far from the deadline should not be very effective for naive decision makers because these decision makers will delay completion of the task, and end up forgetting about the task again over the course of that delay.



Several psychological factors likely contribute to naivete about one's future inattention. One psychological channel is motivational: people like to hold favorable beliefs about the future or about their own traits (Camerer, 1997). A second psychological channel is cognitive: people overuse their current disposition to predict their future disposition, as in projection bias (Loewenstein et al., 2003). A person focusing on a certain task on Tuesday will simply have trouble imagining how he could forget about this task on Wednesday or Thursday.

### **1.2.5 Remarks**

Note that in the baseline model proposed thus far, the agent has no control of the cues in his environment. Indeed, many types of cues are not set by the DM himself: cues coming from interested parties, incidental cues such as conversations with others, or random events (e.g, a new rebate opportunity reminds the DM of a current rebate that is about to expire). This starting point abstracts from the fact that people also use reminder technologies such as calendars, alarm clocks, "to do" lists, etc. In Appendix A.1.2, I consider a more general environment in which the DM first makes additional investments in cues, and then participates in the subgame corresponding to the baseline model defined in this section. In the appendix, I discuss the extent to which investment in cues modifies the results in the paper, and argue that the qualitative results remain largely unchanged. Although a *sophisticated* DM who can cheaply purchase a perfect cue technology and is always attentive to the possibility of this investment option can essentially eliminate his inattention, I argue in appendix A.1.2 that this is an extreme case.

### **1.2.6 Relation to economics work on limited attention**

By studying inattention to *actions*, rather than *information*, my model differs from most existing economic models of inattention. The model in this paper complements models of limited attention in information processing (Sims, 2003; Gabaix and Laibson, 2005; Peng and Xiong, 2006; Reis, 2006; Schwartzstein, 2012; Woodford, 2012; Caplin and Dean, 2013b,a) and models of disproportionate focusing on some attributes over others (Gabaix and

Laibson, 2006; Gabaix, 2012; Köszegi and Szeidl, 2013; Bordalo et al., 2013). The information processing models capture the idea that decision makers may have cognitive limitations in how much information they can use to evaluate an option; my model, in contrast, captures the idea that a decision maker may form a clear intention for how he would like to act in the future, but then fails to follow through on that intention because it is not top of mind. Like the information processing models, the focusing models also lead to biases in the *evaluation* of options, but not to the gap between intentions and future actions captured in this paper. Moreover, neither the information processing models nor the focusing models allow for uninformative cues or past actions to directly influence attentiveness. In Chun et al.'s (2011) terminology, existing economic models would be classified as models of *external attention*—the modulation of how external stimuli or new information is processed—while my model would be classified as a model of *internal attention*—the modulation of internally generated information such as intentions and goals.

Mullainathan (2002) studies limited recall of *information*. The idea that repetition and cues increase attentiveness parallels Mullainathan's modeling of the rehearsal and associativity properties of recall memory. Closely related to this paper, Holman and Zaidi (2010) formulate a model of "prospective memory" ("memory for action") for a decision environment in which there are no external cues and an action can be taken only once (and thus there is no scope for repetition to increase future attentiveness).<sup>16</sup> Their model is formally nested as a special case of the model in this paper. Also related is the model that Karlan et al. (2012a)

---

<sup>16</sup>Ericson (2010) also considers a model in which present-biased agents permanently forget about a task with some probability. "Forgetting," as applied to intentions and actions, might not be the right term, however. In contrast to a person focused on recalling a historical date but consciously failing, the person who absent-mindedly fails to take his medication is not aware of the memory retrieval failure while it is occurring—the intention has been stored in memory and is in tact, but the problem is that the person has simply not directed his attention to retrieving that intention from memory. Some psychologists use the term "prospective forgetting" in laboratory paradigms in which subjects fail to respond to a certain cue or prompt after short time delays. Applied to field behaviors, however, Dismukes (2012) points out that the term "is something of a misnomer, given that what it refers to involves the cognitive process of planning, attention, and task management as much as involves memory. After forming an intention, individuals often become engaged with various ongoing tasks and, in most everyday situations, cannot hold the deferred intention in a focal attention." As footnote 11 also points out, laboratory studies of what cognitive psychologists often call prospective forgetting don't necessarily capture the same mechanisms that operate in field behaviors. I suggest that a better description of the underlying psychology is inattention to previously formed intentions, choices, and actions.

develop for the specific context of consumption and savings with “lumpy expenditures.” Karlan et al. (2012a) assume that consumers might be inattentive to certain aspects of a decision—namely, potential future expenditures—but that cues can direct attention to those aspects of a decision.

## 1.3 Repeated actions

### 1.3.1 Theory

In this section, I investigate the model’s implications for decisions that are made repeatedly and on a regular basis. These include health behaviors such as taking medication or making plans to exercise, actions affecting energy conservation such as turning of energy-using appliances and adjusting the thermostat at peak hours, and actions constituting parts of workplace routines.

Formally, the DM learns about the repeated action in period  $t = 0$  and chooses  $x_t \in \{d, a\}$  in each period  $t = 1, \dots, T$ . The DM’s utility from choosing  $x_t = a$  in period  $t$  is  $b_t + \xi_t$ . Throughout this section, I will rely on the stronger assumptions A3’ and A4’, which state that rehearsal strictly increases next period’s probability of attention, and that this diminishes the impact of cues. I will also assume that  $b_t + \bar{\xi} > 0$  and  $b_t + \underline{\xi} < 0$  for all  $t$ , so that  $Pr(b_t + \bar{\xi} > 0) > 0$  and  $Pr(b_t + \underline{\xi} < 0) > 0$  for all  $t$ . Finally, I will assume that for all  $t \geq 1$ , the DM will be inattentive with positive probability:  $\gamma_t(1, a) < 1$  for all  $t = 1, 2, \dots, T$ .

A naive DM will choose  $x_t = a$  if and only if doing so generates positive flow utility that period:  $b_t + \xi_t \geq 0$ .<sup>17</sup> A sophisticated DM’s strategies are more complicated because he considers how his current actions impact the probability of future attentiveness through behavioral rehearsal. A sophisticated DM will always choose  $x_t = a$  when  $b_t + \xi_t \geq 0$ , but

---

<sup>17</sup>In the model of naivete in the body of the paper, naive DMs don’t consider how performing a behavior affects future attentiveness because they assume that they will be perfectly attentive to the behavior in the future. More generally, however, overestimation and naivete about the effects of behavioral rehearsal don’t need to be as tightly linked. The more general models in Appendix A.1.3 allow for a separation between a general overestimation of attention and a more specific naivete about behavioral rehearsal. All results in this section about naive DMs hold under the narrower assumption that the DM simply does not understand that behavioral rehearsal increases the probability of future attentiveness.

he may also choose  $x_t = a$  when  $b_t + \zeta_t < 0$ .

I begin the analysis in the section by characterizing how changes in payoffs  $b_t$  impact behavior in period  $t$ , as well as in all subsequent and prior periods. I will let  $Pr^s(x_t = a)$  and  $Pr^n(x_t = a)$  denote respective probabilities that sophisticated and naive DMs choose  $x_t = a$  in period  $t$ , from the period 0 perspective.<sup>18</sup>

The assumption that  $F$  is atomless and has a density function ensures that for both sophisticated and naive DMs,  $Pr(x_t = a)$  is differentiable in  $b_{t'}$  for any  $t' \geq 1$  (Lemmas 3 and 4 in appendix A.4.1).

**Proposition 1.** 1. For  $t \leq t'$ ,

$$\frac{\partial}{\partial b_t} Pr^n(x_{t'} = a) > 0 \text{ and } \frac{\partial}{\partial b_t} Pr^s(x_{t'} = a) > 0.$$

2. For all  $1 \leq t'' < t$ ,

$$\frac{\partial}{\partial b_t} Pr^n(x_{t''} = a) = 0 \text{ and } \frac{\partial}{\partial b_t} Pr^s(x_{t''} = a) > 0.$$

3. For all  $1 \leq t'' < t \leq t'$ ,

$$\frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^n(x_{t'} = a) > 0 \text{ and } \frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(x_{t'} = a) > 0.$$

Proposition 1 fully characterizes how changes in payoffs affect the behavior of sophisticated and naive DMs. Part 1 states that increasing the payoff to choosing  $x_t = a$  increases the probability that both a naive and a sophisticated DM will choose  $x_{t'} = a$  for  $t' \geq t$ . The intuition is that increasing  $b_t$  increases the probability that conditional on being attentive, both sophisticated and naive DMs find it worthwhile to choose  $x_t = a$ . Because of rehearsal, this then leads to higher probabilities of being attentive in periods  $t + 1, t + 2, \dots$ , and thus higher likelihoods of choosing  $x = a$  in those periods.

Part 2 states that increasing future payoffs to choosing  $x = a$  also increases a sophisticated DM's motivation to invest in future attentiveness through behavioral rehearsal. Thus for

---

<sup>18</sup>That is, probabilities of choosing  $x_t = a$  when conditioning on the null, period 0 history.

$t'' < t$ , a sophisticated DM's likelihood of choosing  $x_{t''} = a$  is increasing in  $b_t$ .

Part 3 states that for  $t'' < t < t'$ , period  $t''$  and period  $t$  payoffs will have complementary effects on period  $t'$  behavior. The intuition is simple: Part 1 shows that increasing  $b_{t''}$  makes the DM more attentive in all future periods. In particular, this means that the DM will be more attentive in period  $t$ , which will increase his responsiveness to period  $t$  payoffs.

Parts 1 and 2 of Proposition 1 show that the model predicts behavioral patterns similar to what theories of rational, taste-based habit formation predict (Becker and Murphy, 1988). The mechanism is quite different, however. Becker and Murphy (1988) assume that repeated consumption generates a "habit stock" that enters the utility function and creates higher marginal utility from future consumption. Most of the applications of the Becker and Murphy (1988) framework have focused on "bad" habits such as addictions, in which past consumption increases the marginal utility from increasing future consumption, but also lowers the level of utility for any given choice of consumption. The attention model in this paper does not lead to, or apply to such bad habits. However, the model provides an alternative mechanism for the various "good" habits discussed at the beginning of this section. Temporarily increasing the returns to taking some action will increase the likelihood of taking that action is taken during the temporary increase in incentives; but, because of rehearsal, the model predicts that the person will then become more attentive to the action, and thus more likely to take it even when the additional incentives are no longer in place. Evidence for such spillover effects has been documented in the case of energy use (Jesso and Rapson, forthcoming) and exercising (Charness and Gneezy, 2009; Acland and Levy, 2011).<sup>19</sup>

What distinguishes the inattention model from models such as Becker and Murphy (1988) are its predictions about how cues will affect behavior and how they will modulate the spillover effects of incentives. I now turn to formally investigating these effects. Analogous

---

<sup>19</sup>In the case of gym attendance, it is probably less likely that the person literally "forgets" to attend the gym after making plans to do so. Rather, it is more likely that infrequent gym goers probably don't think very often about the possibility of attending the gym, and typically don't think to incorporate it into their daily plans and routines. This is consistent with the findings in Charness and Gneezy (2009) that temporary incentives have the biggest effects on irregular gym goers.

to Proposition 1, I begin by characterizing how changes in the cue distribution  $H_t$  affect behavior in period  $t$ , as well as behavior after period  $t$  and before period  $t$ . Throughout, I will write  $H_t^1 >_{FOSD} H_t^2$  if  $H_t^1$  first order stochastically dominates  $H_t^2$ . When comparing outcomes under two different sequences of cue distributions  $\mathbf{H}^1 = (H_1^1, \dots, H_T^1)$  and  $\mathbf{H}^2 = (H_1^2, \dots, H_T^2)$ , I will index the corresponding outcomes with the superscripts  $H^1$  or  $H^2$ . I will set  $\mu_t^i = \int \sigma dH_t^i$ .

**Proposition 2.** *Consider two sequences of cue distributions  $\mathbf{H}^1 = (H_1^1, \dots, H_T^1)$  and  $\mathbf{H}^2 = (H_1^2, \dots, H_T^2)$  such that for some  $t'$ ,  $H_{t'}^2 >_{FOSD} H_{t'}^1$  but  $H_\tau^1 = H_\tau^2$  for  $\tau \neq t'$ . Then*

1.  $Pr^{n,H^2}(x_t = a) > Pr^{n,H^1}(x_t = a)$  for all  $t \geq t'$
2. When  $t' > 1$ ,  $Pr^{n,H^2}(x_t = a) = Pr^{n,H^1}(x_t = a)$  for all  $t < t'$ .
3.  $Pr^{s,H^2}(x_t = a) > Pr^{s,H^1}(x_t = a)$  for all  $t \geq t'$  when a)  $t' = 1$  or b)  $\mu_{t'}^2$  is sufficiently close to 1.
4. When  $t' > 1$ ,  $Pr^{s,H^2}(x_t = a) < Pr^{s,H^1}(x_t = a)$  for all  $t < t'$ .
5.  $Pr^{n,H^2}(x_t = a) - Pr^{n,H^1}(x_t = a)$  is decreasing  $H_{t''}$  for all  $t'' < t' \leq t$ .

Part 1 states that adding cues unambiguously increase the likelihood that a naive DM chooses  $x_t = a$  both at the time that the cues arrive and in all future periods. By assumption, a period  $t_2$  cue increases the likelihood that the DM is attentive in period  $t_2$ . And because of rehearsal, this effect spills over into all periods following  $t_2$ . Part 2 simply says that naive DMs will not adjust their behavior in anticipation of future cues.

An analog to part 1 holds for sophisticated DMs when the change in cues occurs in period 1, as in part (a) of part 3. Generally, however, the impact on sophisticated DMs' behavior is more nuanced because of another channel through which cues affect behavior: sophisticated DMs realize that the more likely they are to be reminded of the behavior in the future, the less important it is to make it habitual. Thus, the anticipation of future cues crowds out sophisticated DMs' motive to invest in making the behavior more habitual, as

formalized in part 4 of Proposition 2.<sup>20</sup> When the additional cues are sufficiently strong, however, as in condition (b) of part 3 of the proposition, the crowd out effect becomes relatively weak.

Finally, part 5 of Proposition 2 states that for naive DMs, temporally separated cues always have substitutable effects on behavior. The more cues the DM gets in period 1, for example, the smaller the impact of period 6 cues on period  $t \geq 6$  behavior. Intuitively, this is because the more cues the DM gets in period 1, the more likely he is to be attentive in period 6 (by part 1 of the proposition); as a consequence, there is less scope for period 6 cues to make him more attentive that period. This result that temporally separated cues are substitutes is in contrast to part 3 of Proposition 1, which shows that increases in temporally separated *payoffs* have *complementary* effects on behavior.

Recent evidence on exercise (Calzolari and Nardotto, 2012) and energy use (Allcott and Rogers, 2012; Gilbert and Zivin, 2013) is consistent with parts 1 and 3 of Proposition 2, and complements the evidence on habit formation in those domains. Calzolari and Nardotto (2012) show that over the course of 6 months, weekly reminders to attend the gym increase attendance from an average of 8.19 visits per month to an average of 9.31 visits per month. Moreover, they find that the reminder intervention continues to have an effect on behavior for up to 3 months after it ends.

Gilbert and Zivin (2013) show that households reduce consumption by 0.6% to 1% following each electricity bill. Allcott and Rogers (2012) show that following each home energy report, consumers are more likely to engage in repeated actions such as turning off lights, unplugging unused electronics, and adjusting thermostats. Interestingly, Allcott and Rogers (2012) also find that the impact of each additional home energy report diminishes with time. While Allcott and Rogers (2012) propose that one possible reason for this is

---

<sup>20</sup>In the baseline model presented here, sophisticated DMs can increase attentiveness to  $x_t = a$  only through repeated choice of the action. In the more general model in Appendix A.1.2, sophisticated DMs could also increase attentiveness through investments in reminder technologies. Either way, when a sophisticated DM's investment choice is discrete, an additional cue in period  $t_2$  can indirectly lower the probability of this DM's period  $t_2$  attentiveness by crowding out investments in increasing the likelihood of being attentive in the future.

“desensitization” to previously encountered cues,<sup>21</sup> part 5 of Proposition 2 shows that this effect of diminishing marginal returns to additional cues can arise endogenously as a consequence of rehearsal. In the advertising setting examined in Section 1.5.2, I further extend the prediction that temporally separated cues will be substitutes, and examine the implications for optimal policies of reminder advertising.

In addition to their direct impact on behavior, cues also modify how behavior responds to incentives. The next two propositions characterize how 1) period  $t_1$  cues change the response to period  $t_2 \geq t_1$  incentives and how 2) period  $t_1$  incentives change the response to period  $t_2 > t_1$  cues.

**Proposition 3.** *Take  $t_1 \leq t_2$  and consider two sequences of cue distributions  $\mathbf{H}^1 = (H_1^1, \dots, H_T^1)$  and  $\mathbf{H}^2 = (H_1^2, \dots, H_T^2)$  such that for some  $t_1$ ,  $H_{t_1}^2 >_{\text{FOSD}} H_{t_1}^1$ , but  $H_\tau^1 = H_\tau^2$  for  $\tau \neq t_1$ .*

1. For all  $t \geq t_2$ ,

$$\frac{\partial P_{r^n, H^2}(x_t = a)}{\partial b_{t_2}} > \frac{\partial P_{r^n, H^1}(x_t = a)}{\partial b_{t_2}}.$$

2. If  $t_1 = 1$ , then for all  $t \geq 1$ ,

$$\frac{\partial P_{r^s, H^2}(x_t = a)}{\partial b_{t_2}} > \frac{\partial P_{r^s, H^1}(x_t = a)}{\partial b_{t_2}}.$$

Part 1 of Proposition 3 shows that receiving additional cues prior to, or during period  $t$ , makes the naive DM’s behavior more responsive to period  $t$  incentives. In particular, the DM is more responsive in period  $t$  and, because of rehearsal, this heightened responsiveness also spills over into periods  $t + 1, \dots, T$ . Thus, in the specific sense described in Proposition 3, cues *amplify* the effects of incentives.

Intuitively, the DM cannot respond to incentives for choosing  $x_t = a$  unless he is attentive. Thus, when adding cues increases the likelihood of attentiveness, it also magnifies the response to incentives. For a naive DM, this leads to the straightforward prediction that increasing cues in some period  $t'$  magnifies the response to incentives in any period  $t \geq t'$ .

---

<sup>21</sup>See Section 1.6 for a discussion of this effect.



For a sophisticated DM, matters can be more complicated because, by part 4 of Proposition 2, an increase in period  $t$  cues crowds out the motivation to choose  $x_{t'} = a$  in periods  $t' < t$ . In particular, the DM anticipating fewer cues in the future may also be more sensitive to changes in future payoffs due to the strategic rehearsal motive. When  $t_2 = 1$ , however, so that the strategic rehearsal motive is shutdown, the result also holds for sophisticated DMs. I have not yet been able to find more general conditions under which an analog to part 1 of Proposition 3 holds for sophisticated DMs.

The results of Proposition 3 are in line with recent evidence on residential electricity use. Jessoe and Rapson (forthcoming) study how energy use responds to temporary price increases, and find that supplementing the usual price change notifications with additional cues in the form of electricity meters makes people significantly more elastic to price changes. The intuitive explanation provided by Proposition 3 is that the electricity meters increase the likelihood of being attentive to energy use, and thus increase the fraction of people who take actions to reduce energy consumption during a temporarily high price. Appendix A.1.4 derives a formal corollary of Proposition 3 for energy use elasticities.

At the same time, there is also an important sense in which increasing cues can diminish the impact of incentives. In particular, the intertemporal spillover effect of period  $t_1$  incentives becomes less and less pronounced as period  $t_2 > t_1$  cues are increased:

**Proposition 4.** *Take  $t_1 < t_2$  and consider two sequences of cue distributions  $\mathbf{H}^1 = (H_1^1, \dots, H_T^1)$  and  $\mathbf{H}^2 = (H_1^2, \dots, H_T^2)$  such that for some  $t_2$ ,  $H_{t_2}^2 >_{FOSD} H_{t_2}^1$ , but  $H_\tau^1 = H_\tau^2$  for  $\tau \neq t_2$ .*

1. For all  $t \geq t_2$ ,

$$\frac{\partial Pr^{n,H^2}(x_t = a)}{\partial b_{t_1}} < \frac{\partial Pr^{n,H^1}(x_t = a)}{\partial b_{t_1}}.$$

2. For all  $t \geq t_2$ ,

$$\frac{\partial Pr^{s,H^2}(x_t = a)}{\partial b_{t_1}} < \frac{\partial Pr^{s,H^1}(x_t = a)}{\partial b_{t_1}}$$

when  $\mu_{t_2}$  is sufficiently close to 1.

Part 1 of Proposition 4 shows that for a naive DM, the effect of period  $t_1$  incentives on behavior in periods  $t \geq t_2$  becomes less and less pronounced as period  $t_2$  cues are increased.

Part 2 establishes the same result for a sophisticated DM. The statement is weaker because the period  $t_2$  cues can change how a sophisticated DM responds to incentives in period  $t_1 < t_2$ . In particular, it is possible that the sophisticated DM may become more responsive to period  $t_1$  incentives as period  $t_2$  cues are increased. Nonetheless, a sufficiently large increase in period  $t_2$  cues will diminish the impact of period  $t_1$  incentives even for a sophisticated DM. Intuitively, if period  $t_1$  incentives impact future behavior by changing the likelihood of future attentiveness, then their effects will become negligible if future cues are sufficiently strong to guarantee almost perfect attentiveness.

Proposition 4 is particularly helpful for distinguishing taste-based theories of habit formation from the theory proposed in this paper. The simple but diagnostic prediction that reminders should have the largest effect when the behavior has not been recently performed is difficult to rationalize with any other existing theories, or even combinations of existing models. For example, if people had habit-forming preferences *and* were inattentive with some probability  $p$  that was independent of their past behavior (but could be bolstered with cues), then they would behave in a manner opposite to Proposition 4.<sup>22</sup>

Yet to my knowledge, this interaction between repetition of a behavior and subsequent reminders has not been investigated in economics, psychology, or related disciplines. Section 1.3.2 below reports an experiment investigating this relationship.

## 1.3.2 Experimental Evidence

### Design and Procedures

To examine the interaction between action repetition and reminders, I conducted an online, real-effort experiment spanning three weeks of daily tasks. The daily task that subjects had to perform repeatedly over the course of 3 weeks was to complete a short, online daily survey. The repetitive nature of this short, simple task closely resembles taking daily

---

<sup>22</sup>Intuitively, this is because for these types of decision makers, a period  $t$  reminder would only amplify the impact that past behavior would have on the preference for choosing  $x_t = a$ . The logic behind this comparative static is analogous to the logic behind Proposition 3.

medication or online browsing behaviors. Disguising this task as a “survey study” and not disclosing the true purpose of the experiment made it possible to study how subjects *naturally* approach these kinds of real-world behaviors.<sup>23</sup>

The  $2 \times 2$  factorial design varied whether or not subjects were given the opportunity to complete the task in the middle week, and whether or not subjects received reminders in the third week. The experiment consisted of three phases, depicted in Figure 1.1:

Registration Phase (Day 0) Subjects interested in participating in the experiment were directed to the study website, where they created an account for the experiment. Upon creating an account, they were randomized into the study conditions described below. They then received a short overview of the 3-week experiment and completed 5 minutes of demographic and lifestyle questions for which they received \$3. They were then given detailed instructions explaining their “daily survey” task, which started the *day after* they signed up and spanned the next three weeks. Immediately upon completing the registration phase, subjects received an automatic email with a copy of their electronic consent form, and all study instructions (which included a link to the study website).

Daily Survey Phase (Days 1-21) The simple task subjects faced in days 1-21 was to log into the study website (using the email and password they registered for this experiment), and to report their current level of excitement, happiness, stress, and worry using a 1-5 Likert scale. Subjects received \$1 for each day that they completed this daily survey. To ensure a time break between any two consecutive survey completions, the daily survey was not available between 12:00 a.m. and 4:59 a.m. each day. In all four conditions, subjects did not initially know if the survey would be available to them in week 2 of the study. They were told that it would be available in week 2 with 50% chance, and that they would be notified of this at 5:00 a.m. on the first day of week 2. The four conditions were as follows:

- *No interruption / No reminders.* Subjects in this condition received only 2 communica-

---

<sup>23</sup>This experiment involved absolutely no deception, however—only *incomplete disclosure*. Subjects were told all details of the experiment other than the true purpose of the study (and were not led to believe that the purpose of the study was anything other than what it is).

tions throughout the study:<sup>24</sup> An email on day 1 reminding them to start completing the daily survey, and an email on day 8 informing them that the survey would be available for the next 14 days.

- *Week 2 interruption / No reminders.* This condition was identical to the condition above, except that on day 8, these subjects received email notification that the daily survey would not be available to them for the next 7 days, but that it would be made available again for the last 7 days of the study.
- *No interruption / Week 3 reminders.* This condition was identical to the *No interruption / No reminders* condition, except that subjects received an email each day of week 3 stating “This is a reminder that the daily survey is available to you between 5:00 a.m. and 11:59 p.m. today.”
- *Week 2 interruption / Week 3 reminders.* This condition was identical to the *Week 2 interruption / No reminders* condition, except that subjects received an email each day of week 3 stating “This is a reminder that the daily survey is available to you between 5:00 a.m. and 11:59 p.m. today.”

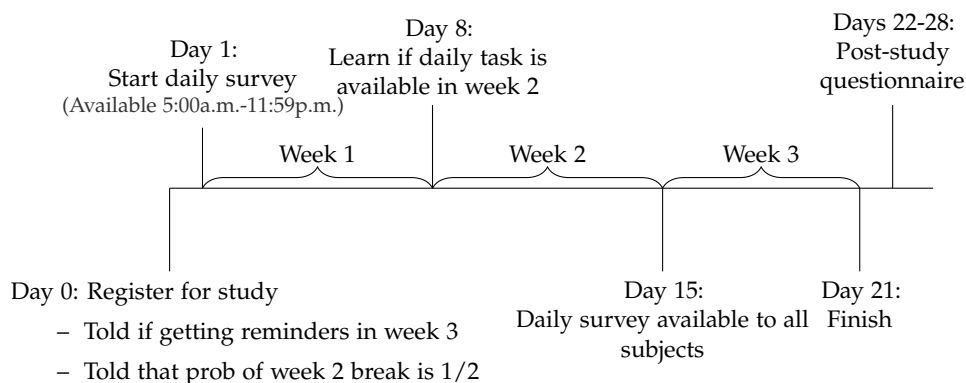
Post-task phase At the end of the three weeks of the daily survey, subjects completed a short closing questionnaire, for which they were paid an additional \$3.<sup>25</sup> The closing questions asked about subjects’ use of reminder devices to complete the daily survey, as well as subjects’ daily routines (if any) for completing the daily survey. The questions are described and analyzed in more detail below, as well as in Appendix A.2.

In all conditions, subjects were informed on day 0 of all of the communications they would receive, and were told that they would receive no other communications.

---

<sup>24</sup>All communications were emailed out at 5:00 a.m.

<sup>25</sup>Subjects who completed the daily survey on day 21 were immediately prompted to complete the 5 minute closing questionnaire. Subjects who did not complete the daily survey on day 21 were sent an email the next day inviting them to complete the closing questionnaire. All subjects had one week to complete the closing questionnaire, and were sent a reminder each day of that week to complete the questionnaire. If subjects did not complete the closing questions within one week, they did not receive the additional \$3, and were emailed a receipt for their other earnings.



**Figure 1.1:** *Timeline for Repeated Action Experiment*

All earnings were emailed to subjects in the form of an Amazon.com gift certificate after the completion of the study.<sup>26</sup>

Formally, the week 2 interruption can be modeled as creating substantially lower benefits from performing the action.<sup>27</sup> Notice that in contrast to the decisions studied in the theory section, however, whether or not the week-long interruption would occur was not known with certainty. A consequence of this feature is that conditional on the reminders treatment, week 1 behavior should be identical for subjects with a week 2 break and without a week 2 break. Controlling for week 1 behavior thus generates additional statistical power. Propositions 24 and 25 in Appendix A.2 generalize Propositions 1 and 4 to the experimental setting studied here.

---

<sup>26</sup>For subjects who completed the closing survey and signed a study receipt, these payments were emailed within 24 hours. Subjects who never completed the closing survey were emailed a link to a receipt for their previous earnings 8 days after the end of the daily survey phase.

<sup>27</sup>So technically, participants very serious about keeping this activity on their mind can always log into the study site, and then pretend to answer the 4 feelings question. Alternatively, this week 2 break can also be modeled as creating such low benefits that no subject would ever want to complete the survey in week 2. A few subjects did log in to the study site on the first day of week 2, but this probably reflects confusion.

## Results

The experiment was run through the Harvard Decision Science Laboratory (HDSL) during July and August 2013, and enrollment was limited to members of the HDSL subject pool.<sup>28</sup> A total of 187 subjects completed the Registration phase. Of these 187, 7 subjects did not complete any daily surveys during the first week, and are thus excluded from all subsequent analysis. This exclusion is particularly reasonable because these 7 subjects also did not complete any surveys during the subsequent weeks, meaning that in my sample, not doing the survey in week 1 was perfectly predictive of dropping out of the experiment altogether. Recruitment took place over the course of 6 days: on July 15,16,17 subjects who signed up were randomized into the no reminders conditions (105/180 subjects), while on August 6,7,8 subjects who signed up were randomized into the reminders conditions (75/180) subjects.<sup>29</sup>

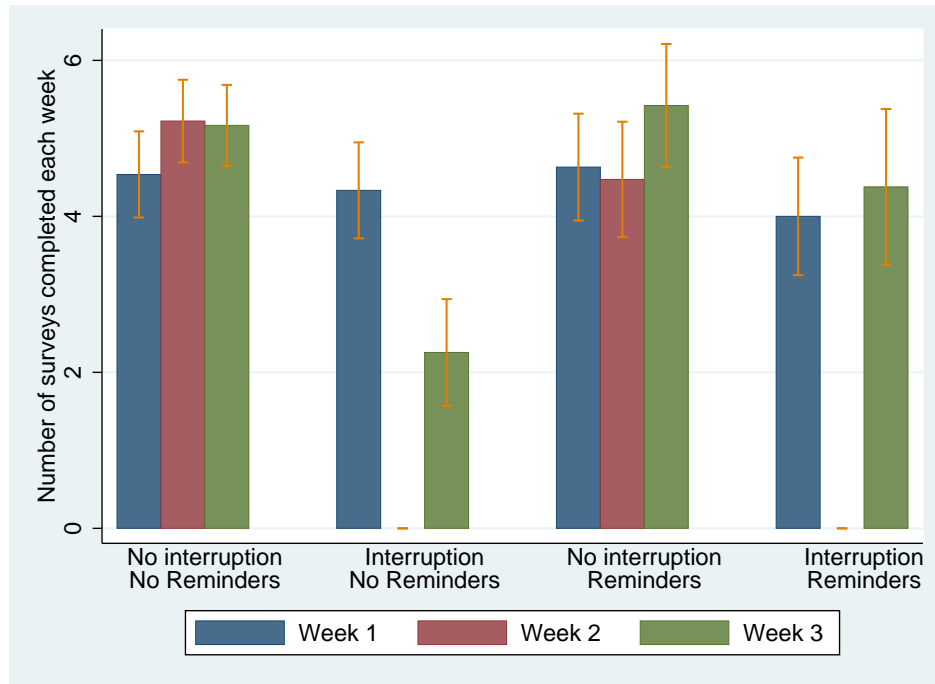
Figure 1.2 shows weekly averages for subjects in all 4 conditions, with error bars corresponding to 95% confidence intervals. Strikingly consistent with the habit effect identified in Proposition 1 (and extended to this experimental design in Proposition 25) subjects who did not receive reminders in week 3 completed significantly fewer surveys if they experienced a week 3 interruption. While uninterrupted subjects completed an average of 5.2 out of 7 surveys in week 3, subjects who experienced a week 2 interruption completed an average of only 2.3 surveys in week 3.

At the same time, the effect of the interruption is almost completely undone by the daily reminders in week 3. While week 3 reminders have almost no effect on uninterrupted subjects (a minor increase of 0.3 surveys), they increase the interrupted subjects' week 3 completion rates from an average of 2.3 to an average of 4.4 surveys.

---

<sup>28</sup>All members of HDSL were eligible to participate in this experiment. After signing up through HDSL's SONA system, subjects were given a link to the study site, where they could initiate the registration process.

<sup>29</sup>That the reminders and no reminders conditions were split up across time might seem worrisome because of time-dependent shocks. However, even if calendar date random effects did cause differences in levels, it is hard to think of a reason for why they should confound the estimate of the interaction effect between the week 2 interruption and week 3 reminders. But to address the concern that calendar date random effects might somehow be inflating the interaction effect between week 2 interruption and reminders, I show that the results are also robust to clustering at both the subject level and calendar date level. I also show that subjects' demographics do not vary across the four experimental conditions, and that all results are robust to demographic controls.



**Figure 1.2:** *Surveys Completed Per Week, by Experimental Condition*

Table 1.1 quantifies the effects in Figure 1.2 by estimating linear probability models of daily survey completion, with robust standard errors clustered at the subject level.<sup>30</sup> Specification (1) includes only the experimental conditions as the independent variables, while specifications (2) and (3) also include the average daily completion rate in week 1 (i.e., number of surveys completed in week 1 divided by 7) as a covariate. All three specifications show that a break between weeks 1 and 3 reduces the probability that a subject completes a survey on any given day of week 3 by about 35 percentage points. At the same time, the regressions show that most of this effect is mitigated with daily reminders in week 3: the interaction effect between experiencing a week 2 break and receiving week 3 reminders is about 25 percentage points, which is significant at the 5% level in specification (1) and significant at the 1% level in specifications (2) and (3).

In the first specification, the effect of the week 2 interruption on week 3 performance in

---

<sup>30</sup>Table A.3 shows that the results are unchanged with robust two-way clustering at both the subject level and the calendar date level.

**Table 1.1:** *Probability of Completing a Survey on any Given Day in Week 3*

Pr(complete survey)	(1)	(2)	(3)
Interruption	-0.416*** (0.061)	-0.398*** (0.054)	-0.401*** (0.054)
Reminders	0.036 (0.067)	0.028 (0.051)	-0.121 (0.137)
Interruption*Reminders	0.267** (0.108)	0.305*** (0.089)	0.320*** (0.091)
Week1Avg		0.625*** (0.078)	0.527*** (0.093)
Week1Avg*Reminders			0.226 (0.162)
Adj. $R^2$	0.139	0.289	0.293
Observations	1260	1260	1260

*Notes:* This table estimates a linear probability model of completing the daily survey in week 3 of the study. The variable “Interruption” equals 1 if the daily survey was available in week 3 and equals 0 otherwise. The variable “Reminders” equals 1 if subject received daily reminders in week 3 and equals 0 otherwise. The variable “Week1avg” denotes the fraction of surveys completed in week 1. Robust standard errors clustered at subject level. All regressions include controls for day of week and day in study. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

the reminders conditions is still a marginally significant 13.9 percentage points ( $p = 0.097$ ), but some of that is due to a slightly unbalanced randomization. As shown in Figure 1.2, uninterrupted subjects also had a slightly higher week 1 average. Specification (2) controls for the week 1 average under the assumption that a subject’s week 1 behavior should not be affected by study condition. This is the right assumption for naive subjects who do not adjust their behavior in anticipation of future cues, but not the right assumption for sophisticated subjects. Specification (3) addresses this theoretical point by interacting the week 1 average with whether or not a subject was in a reminders condition. In both of these specifications, the impact of the interruption is estimated to be 8 to 9 percentage points, which is not statistically different from zero ( $p = 0.205$  and  $p = 0.274$ , respectively). Appendix A.2 shows that all of these results are unchanged when controlling for various demographic characteristics (Table A.2), and when calendar date random effects are taken into account (Table A.3).

The end of study questionnaire, which was completed by 172 of the 180 subjects, provides



further evidence in support of the model.<sup>31</sup> In response to the question “On a scale of 1-5, with 1 being ‘not at all’ and 5 being ‘very much,’ to what extent did the following factors contribute to you not filling out the daily surveys” subjects gave an average rating of 3.9 for “Forgot”, 1.8 for “Didn’t have time,” 1.16 for “Didn’t feel like it,” 1.64 for “Could not remember log in information,” and 1.58 for “No internet access.” The difference between “Forgot” and each of the other reasons is significant at  $p < 0.001$  in pairwise t-tests and signed rank tests.

During the registration phase, subjects were also asked to predict how many times they would complete the daily survey each week. They were asked four questions: 1) how many times they would complete the survey in week 1; 2) how many times they would complete the survey in week 2 if it was available then; 3) how many times they would complete the survey in week 3 if it was available in week 2; 4) how many times they would complete the survey in week 3 if it was not available in week 2. Although unincentivized and thus needing to be interpreted with caution, subjects’ day 0 forecasts of their daily survey completion rates support the naive model. Of the 168/180 subjects who completed the predictions questions and gave answers within range,<sup>32</sup> the average forecast was between 6.72 and 6.9 surveys for each week. Subjects did not predict that having a week 2 interruption would affect their week 3 behavior, and their forecasts were insensitive to whether or not they were in the reminders or no reminders conditions. For all weeks and all conditions, the forecasts were significantly higher than actual behavior (paired t-tests  $p < 0.001$  in all comparisons).

Of course, some of the initial confidence may reflect subjects’ intentions to use reminder technologies, which, as analyzed in more detail in Appendix A.2, were employed by 30% of the subjects. There is, however, no statistically significant interaction between forecasts and using reminder technologies. Thus the 70% of the subjects who used no reminder

---

<sup>31</sup>It was completed by 102/105 subjects in the no reminders conditions, and by 71/75 subjects in the reminders conditions. Subjects were given one week to complete the end of study questionnaire, and received a daily reminder each day of the week to complete it.

<sup>32</sup>98/105 in the no reminders conditions and 74/75 in the reminders conditions completed the questions. Of these, 2/98 subjects in the no reminders conditions and 2/74 in the reminders conditions gave out of range answers.

technologies appear to be quite naive. Appendix A.2 also shows that subjects' week 3 performance was less affected by the week 2 interruption if they reported using a reminder technology such as a calendar (typically Google calendar), writing notes or asking others to remind them.<sup>33</sup> Moreover, these subjects were also less affected by the week 3 reminders.

## 1.4 Tasks With Deadlines

### 1.4.1 Theory

In this section I investigate the model's basic implications for decisions involving a single task that must be completed by a deadline  $T$ . The task might be making a savings deposit, applying for a rebate, paying a bill, canceling a subscription, scheduling an appointment, and so forth.

The DM learns about the task in period  $t = 0$ , and must complete the task at some time between period  $t = 1$  and period  $t = T$ . The task can be completed only once, with  $x_t = a$  corresponding to completing the task in period  $t$ , and with  $x_t = d$  corresponding to not doing the task in period  $t$ . The DM obtains payoff  $b_t + \zeta_t$  if he completes the task in period  $t$ , and obtains a payoff of 0 if he never completes it. Throughout this section, I will consider a DM who is always attentive in period 0 when he first learns about the task:  $\gamma_0 = 1$ .

Throughout much of the analysis in this section, I will assume that the distribution  $F$ , deterministic payoffs  $\mathbf{b} = (b_1, b_2, \dots)$ , and cue distributions  $\mathbf{H} = (H_1, H_2, \dots)$  are fixed, and that the choice of deadline  $T$  is applied to the (infinite) sequences  $\mathbf{b}$  and  $\mathbf{H}$ , specifying that  $X_t = \{d\}$  for all  $t > T$ . Thus for any two deadlines  $T_1 < T_2$ , a DM given a deadline  $T_2$  faces the same payoff and cue distributions for the first  $T_1$  periods as the DM facing a deadline  $T_1$ .

I assume that for an *inattentive* DM there is some  $\bar{z} < 1$  such that  $\gamma_t(1, d) \leq \bar{z}$  for all  $t \geq 1$ . A perfectly attentive DM will be defined as having  $\gamma_t(1, d) = 1$  for all  $t$ .

The next proposition characterizes how inattention affects the probability of completing a task by some time  $t \leq T$ . In stating the results, let  $Q^s$  denote the random variable

---

<sup>33</sup>Of course, standard caveats about endogeneity apply.

corresponding to the period in which a sophisticated DM completes the task, with  $Q^s = \infty$  if the task is never completed. Let  $Q^n$  and  $Q^{pa}$  be defined analogously for naive DMs and for perfectly attentive DMs, respectively.

**Proposition 5.** 1.  $Pr(Q^n \leq t) < Pr(Q^{pa} \leq t)$  for all  $t \leq T$ .

2. Suppose that  $b_t = b$  for all  $t$  and that  $b + \underline{\xi} \geq 0$ . Then for each  $t^\dagger \geq 1$ , sequence of cue distributions  $\mathbf{H}$ , and atomless  $F$ , there is a sufficiently large  $T$  such that  $Pr(Q^s \leq t) > Pr(Q^{pa} \leq t)$  for all  $t \leq t^\dagger$ , but such that  $Pr(Q^s \leq T) < Pr(Q^{pa} \leq T)$ .

3. For each  $t \geq 1$ , there exists a  $z^\dagger > 0$  such that  $Pr(Q^s \leq t) < Pr(Q^{pa} \leq t)$  whenever  $\gamma_{t'}(1, d) \leq z^\dagger$  for all  $t' \leq t$ .

For naive DMs, the effect of inattention is straightforward: for any  $t \leq T$ , they are less likely to complete the task by period  $t$  than perfectly attentive DMs. This is because conditional on being attentive, they follow the same strategy as perfectly attentive DMs, but they never complete the task when inattentive. Part 1 of Proposition 5 formalizes this idea.

The behavior of inattentive but sophisticated DMs, however, is less straightforward to analyze because they may follow the strategy of “I should do it while it’s on my mind.” Part 2 of Proposition 5 shows that for sufficiently long deadlines and sufficiently high value tasks, inattentive but sophisticated DMs will actually be more likely to complete the task by some early date than perfectly attentive DMs. The intuition is that as the deadline becomes longer, the option value to waiting increases for perfectly attentive DMs, making it more and more attractive to put off the task until later. Inattentive but sophisticated DMs, however, will be afraid to put off the high value task to a later time because of the possibility of forgetting about it.

The last part of Proposition 5 is a partial analogue to the result for naive DMs. It is only a partial analogue because even for  $t = T$ , inattentive but sophisticated DMs may be more likely to complete the task due to the fact that they are less likely to put it off than perfectly attentive DMs. Part 3 shows, however, that if the sophisticated DM is sufficiently inattentive, then he will be less likely than the naive DM to complete the task by any period  $t \leq T$ .

I now turn to the question of how deadline length affects welfare and completion rates. In stating the formal results below, I will let  $Q^{s,T_i}$ ,  $Q^{n,T_i}$ ,  $Q^{pa,T_i}$  denote the random variables corresponding to when sophisticated, naive, and perfectly attentive DMs, respectively, complete the task (with  $Q^{T_i} = \infty$  if the task is never completed) when facing deadline  $T_i$ . I will let  $V_0^{s,T_i}$ ,  $V_0^{n,T_i}$ ,  $V_0^{pa,T_i}$  denote the ex-ante expected utilities of sophisticated, naive, and perfectly attentive DMs, respectively, when facing a deadline  $T_i$ .

While many of the results will be stated for any sequence of payoffs  $\mathbf{b}$ , a leading case that will be considered throughout this section is the case of time-invariant payoffs:  $b_t = b$  for all  $t \leq T$ . Most of the examples and applications discussed in this section involve a task that pays a fixed benefit  $b$  but involves a stochastic effort or opportunity cost  $\xi$ . A more general condition that subsumes this special case and that will be used in some of the results is the following:

**Condition I( $T_1, T_2$ )** Given two deadlines  $T_1$  and  $T_2 = T_1 + \Delta > T_1$ : i)  $b_t \leq b_{t+\Delta}$  for all  $t \leq T$  and ii)  $b_t - b_{t-1} \geq b_{t+\Delta} - b_{t-1+\Delta}$  for all  $t \leq T_1$ .

In words, Condition I( $T_1, T_2$ ) states that the last  $T_1$  periods of the longer deadline have a payoff profile that is no lower and no “steeper” than the payoff profile of the shorter deadline  $T_1$ .

Proposition 6 states results for perfectly attentive and for sophisticated DMs.

**Proposition 6.** *Consider two deadlines  $T_1 < T_2$ . Then*

1.  $V_0^{s,T_1} < V_0^{s,T_2}$  and  $V_0^{pa,T_1} < V_0^{pa,T_2}$
2. *Suppose that either Condition I( $T_1, T_2$ ) holds or that  $b_{T_2} + \underline{\xi} \geq 0$ . Then  $\Pr(Q^{pa,T_1} \leq T_1) \leq \Pr(Q^{pa,T_2} \leq T_2)$ .*
3. *Suppose  $b_t = b$  for all  $t \leq T_2$  and that  $\gamma_t(0, d) \geq \underline{z} > 0$  for all  $t$ . Then  $\Pr(Q^{s,T_1} \leq T_1) < \Pr(Q^{s,T_2} \leq T_2)$  for sufficiently high  $b$ .*

Part 1 of Proposition 6 states that a sophisticated or fully attentive DM can only be made better off by longer deadlines. The straightforward intuition is that when given the longer

deadline  $T_2$ , in the first  $T_1$  periods the DM can still behave the same way he would behave if he was given the shorter deadline  $T_1$ .

Part 2 of Proposition 6 analyzes completion rates of DMs who are fully attentive. While longer deadlines unambiguously increase welfare of sophisticated DMs, the effect of deadlines on completion rates is less clear cut in certain special cases. When the task is not very pleasant or important *and* the distribution of task completion payoffs decreases with time, longer deadlines may lead to lower completion rates even for perfectly attentive DMs. As a simple example, suppose that  $\xi \sim U[-1,0]$  and that  $b_1 = 1$  while  $b_2 = 2/3$ . Then when  $T = 1$ , the probability of task completion is 1 because  $b + \xi_1$  is always positive. When  $T = 2$ , the option value of delaying to period 2 makes it suboptimal to complete the task in period 1 whenever  $\xi_1$  is sufficiently close to zero. In period 2, however, the DM completes the task with probability less than 1. This then implies that the overall probability of task completion is less than 1 with a two period deadline.

Part 3 of Proposition 6 proves a result analogous to part 2 for inattentive but sophisticated DMs. When the task is sufficiently important to complete, sophisticated DMs will use the extra time to increase the probability of completing the task.

For naive DMs, however, I show below that extending a deadline can both reduce their welfare and lead to lower completion rates, even in situations in which longer deadlines lead to higher completion rates for sophisticated DMs.

**Proposition 7.** *Consider two deadlines  $T_1 < T_2$ . Then for any  $\mathbf{b}$  and  $F$ , there exists a  $\lambda > 0$  such that  $V_0^{n,T_1} > V_0^{n,T_2}$  and  $Pr(Q^{n,T_1} \leq T_1) > Pr(Q^{n,T_2} \leq T_2)$  if:*

- i)  $\gamma_t(1, d) \geq \gamma_{t+1}(1, d)$  and  $\gamma_t(0, d) \geq \gamma_{t+1}(0, d)$  for all  $t \geq 1$
- ii)  $\gamma_t(1, d) \leq \lambda$  for all  $t > T_1$
- iii)  $\gamma_t(0, d) \leq \lambda \gamma_1(1, d)$  for all  $t > T_1$

Proposition 7 shows that giving a naive DM more time to complete the task can make him worse off and less likely to complete the task if his probability of being attentive to the task decays sufficiently quickly over time, as guaranteed by the three conditions in the

proposition. The simple intuition is that the longer the deadline, the higher is a naive DM's *perceived* option value of delaying the task and waiting for a more opportune time. What a naive DM doesn't realize, however, is that at a later date the task will probably not be on his mind.

Proposition 7 generalizes an insight developed by Holman and Zaidi (2010) for a setting that is formally equivalent to the case of no-cues, time-invariant payoff distributions, and an exponential attention decay curve that limits to zero. The intuition behind conditions (i)-(iii) in the proposition is as follows. Conditions (i) and (ii) ensure that if the DM is attentive in period  $t = 1$ , he will be attentive in period  $t > T_1$  with low probability. Condition (iii) ensures that if the DM is not attentive in period  $t = 1$ , then his probability of being attentive in periods  $t > T_1$  is low *relative* to his probability of being attentive in period  $t = 1$ . Combined, these three conditions ensure that the probability of being attentive in periods  $t > T_1$  is substantially lower than the probability of being attentive in periods  $t \leq T_1$ . The naive DM's perception of periods'  $t > T_1$  is substantially inflated, however. Consequently, he is much less likely to complete the task before period  $T_1$  when given more time, while simultaneously not being very likely to complete the task after period  $T_1$ .

Importantly, the potentially perverse effect of longer deadlines is not generated simply by limited attention, but by the *decay* of attention over time. In fact, with time-invariant payoffs and a constant probability of attention ( $\gamma_t(1, d) = \gamma_t(0, d) = \rho$  for all  $t \geq 1$ ), longer deadlines would always leave a naive DM better off and lead to higher completion probabilities.

The conditions in Proposition 7 also provide insight into the mechanisms that lead to the decay of attention. When cue distributions don't change over time— $H_t = H$  for all  $t \geq 1$ —rehearsal always leads to decay of attention. To see this, suppose that for some  $\bar{\rho} > \underline{\rho}$ ,  $\gamma_t(1, d) = \bar{\rho}$  and  $\gamma_t(0, d) = \underline{\rho}$  for all  $t \geq 1$ . Then the period  $t$  unconditional probabilities of being attentive,  $p_t$ , satisfy  $p_{t+1} = p_t(\bar{\rho} - \underline{\rho}) + \underline{\rho}$ . When  $p_0 = 1$ , the sequence  $\{p_t\}$  can be shown to be decreasing; and Proposition 7 implies that there exists some  $\lambda > 0$  such that  $T_2$  leads to lower welfare and completion rates when  $\bar{\rho} < \lambda$  and  $\underline{\rho} < \lambda\bar{\rho}$ .

Additionally, however, attention decay could also be a consequence of decreasing cues.

Various paperwork lying on the table, conversations with others, or emails at the top of the inbox are examples of cues that are likely to be present when the DM first learns of the task but that will dissipate with time. Even in the extreme case in which  $\gamma_t$  is not affected by mental rehearsal, the cue distributions could still be such that the probability of being attentive in period  $t \geq 1$  is given by, e.g.,  $\gamma_t(1, d) = \gamma_t(0, d) = k\bar{\rho}^t + \underline{\rho}$ . Proposition 7 then again implies that completion rates will be lower under the longer deadline  $T_2$  when  $\bar{\rho}$  is low and when  $\underline{\rho}$  is low *relative* to  $\bar{\rho}$ .

Empirically, an inverse relationship between deadline length and completion rates was first documented by Shafir and Tversky (1992), and subsequently replicated and extended by Shu and Gneezy (2010), Silk (2004), and Janakiraman and Ordóñez (2012). Shafir and Tversky (1992) offered students \$5 to complete a long questionnaire by a given date. One group ( $N = 56$ ) was given a 5 days to complete the questionnaire while the other group ( $N = 58$ ) was given 3 weeks. The rates of return were 60% for the short deadline group and 42% for the long deadline group.<sup>34</sup> Silk (2004) replicated this deadline effect in a setting closely resembling consumer rebates. Shu and Gneezy (2010) replicated this effect with gift certificates for “immediately enjoyable experiences,” such as pastries and movies. And Janakiraman and Ordóñez (2012) replicated this effect in a setting closely resembling product returns.<sup>35</sup> Further supporting naive inattention, Silk (2004) and Shu and Gneezy (2010) also elicited beliefs, and found that while longer deadlines led to lower completion rates, participants predicted higher completion rates with a longer deadline.

The model provides a formal explanation for the evidence above, but also makes

---

<sup>34</sup>Shafir and Tversky (1992) also had a “no definite deadline condition,” but did not provide details on how they assessed completion rates in that condition. Without a definite deadline, even perfectly sophisticated and attentive agents might generate low completion rates by any finite date.

<sup>35</sup>Bertrand et al. (2010) examine how interest rate reductions that last 2, 4, or 6 weeks affect the demand for small, short-term loans by the working poor population in South Africa. Bertrand et al. (2010) find that consumers are significantly more likely to take out a loan with a longer deadline. This result, however, is not in conflict with the experimental evidence discussed in this paragraph. The reason is that a consumer is, mechanically, more likely to generate a need for a loan over the course of 6 weeks rather than over the course of 2 weeks. Conditional on generating a need within 2 weeks, consumers may be less likely to take out a loan from the lender when the interest perk lasts 6 weeks rather than 2; however, this effect is trumped by the fact that far fewer people will generate a need over the course of only two weeks.

predictions about when longer deadlines *should not* lead to lower completion rates. The first prediction is that when the probability of being attentive is bounded away from zero—either because there is always a small chance of encountering a cue, or because  $g(0, d, \sigma)$  is intrinsically bounded away from zero for all  $\sigma$ —the probability of task completion will approach 1 as deadlines become very long. Longer deadlines can unambiguously decrease completion rates only when  $\gamma_t(0, d) = 0$  for all  $t$ —so that the probability of being attentive decays rapidly toward zero.<sup>36</sup>

**Proposition 8.** *Suppose that there is some  $b^* > -\bar{\xi}$  such that  $b_t \rightarrow b^*$ .*

1. *If there is some  $\underline{z} > 0$  such that  $\gamma_t(0, d) \geq \underline{z}$  for all  $t$ , then  $\lim_{T \rightarrow \infty} Pr(Q^{s, T} \leq T) = \lim_{T \rightarrow \infty} Pr(Q^{n, T} \leq T) = 1$ .*
2. *If  $b_t = b$  and if  $\gamma_t(0, d) = 0$  for all  $t$ , then  $\lim_{T \rightarrow \infty} Pr(Q^{n, T} \leq T) = 0$ .*

Combined, Propositions 7 and 8 imply that completion rates will be lowest for “intermediate length” deadlines when cues are rare and/or weak. Figure 1.3 illustrates this “U-shaped” pattern. In both panels  $b_t + \xi_t \sim U[0, 1]$  and the unconditional probability of  $\alpha_t = 1$  is given by  $0.8(0.75)^t + 0.2$ . In the left panel, this decay is generated purely through the mental rehearsal property, while in the right panel this decay is generated purely by changes in cues.<sup>37</sup>

Unfortunately, Proposition 8 is difficult to test because it does not provide guidance on *when* the non-monotonicity should occur, if it does. A more diagnostic and easily testable prediction of the model is that even relatively small changes in cues can alter the effects of a

---

<sup>36</sup>The convergence to zero completion probability is a knife-edge result, however, because if the DM discounted future payoffs by some  $\delta < 1$ , then his probability of task completion would be bounded away from zero. Still, longer deadlines could keep decreasing the completion probability even when  $\delta < 1$ .

<sup>37</sup>There are two caveats to this prediction, however. First, even if  $\gamma_t(0, d)$  is bounded away from zero, then the possibility that the DM might, for example, lose the necessary paperwork to mail in his rebate would, effectively, lead to a task completion decay curve that limits to zero. Second, if the DM is sufficiently present-biased (Laibson, 1997; O’Donoghue and Rabin, 1999) and naive about the present bias, and if the variation in  $\xi$  is sufficiently small, then he will always put off the task until the last period. Combined with decaying, this would lead longer deadlines to unambiguously decrease completion rates. (Relatedly, Ericson (2010) considers a model in which present-biased agents permanently forget about the task with some probability  $\rho$  each period. Part 2 of Proposition 8 shows that even without present bias, longer deadlines can be uniformly bad in such a permanent forgetting model.)



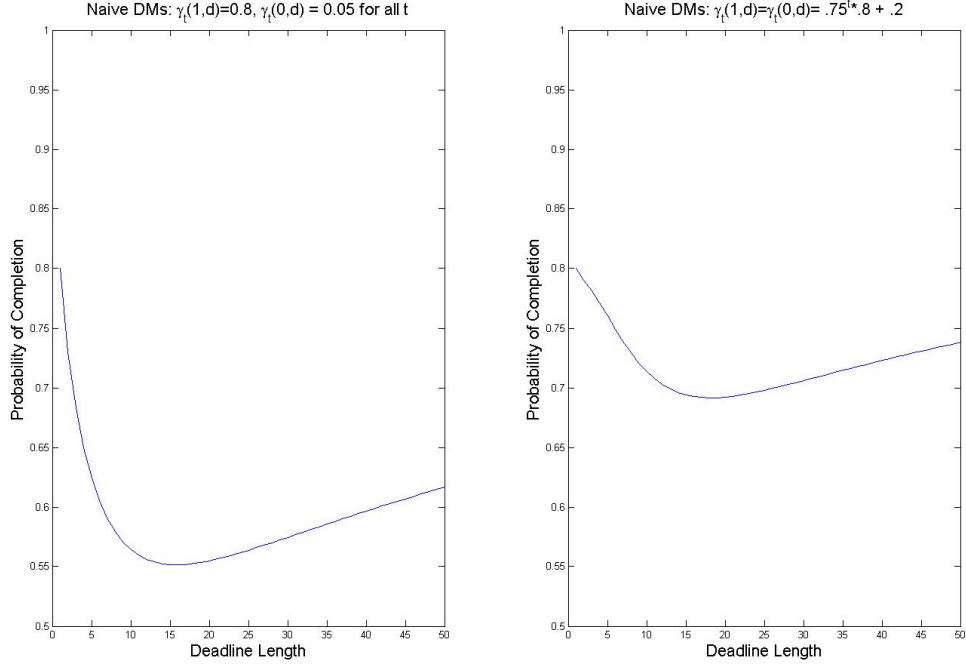


Figure 1.3: “U-shaped” Curve

longer deadline. The rough idea is that *if* longer deadlines decrease task completion due to the decay of attention over time, *then* increasing cues several periods before the deadline to stop the decay can have a very large effect on the completion probability of a DM facing the longer deadline. When attention decay is substantial enough to lead to perverse deadline effects, a DM who faces a long deadline  $T_2$  and has not yet completed the task by period  $T_2 - 1$  will respond very strongly to an increase in period  $T_2 - 1$  cues; this DM simply won't be attentive unless he gets a cue.

To formally state comparative statics about adding cues under two different deadlines  $T_1$  and  $T_2 = T_1 + \Delta > T_1$ , I will begin with *initial* attention probabilities  $\gamma_t^0(\alpha, x)$  and consider increasing them to  $\gamma_t^{\kappa_i}(\alpha, x) = \gamma_t^0(\alpha, x) + (1 - \gamma_t^0(\alpha, x))\kappa_t^i$ , where  $i \in \{1, 2\}$  denotes the resulting probability for each of two deadlines  $T_1$  and  $T_2$ . Note that as in previous analysis, I assume that the cue distributions in periods  $t \leq T_1$  are initially identical in both of the deadlines conditions. But in contrast to the previous analysis, I now allow the modified cue

distributions to differ between deadlines. This set-up makes it possible to compare how, for example, sending a reminder on the last or second-to-last period affects completion rates under short versus long deadlines.

The parametrization of  $\gamma_t^{\kappa_t^i}(\alpha, x)$  is natural and fits into the simple model of cues proposed in the discussion surrounding equation (1.1): Suppose that the DM is initially inattentive with probability  $1 - \gamma$ . Now consider sending reminders whose effect is that a DM who would have otherwise been inattentive is now attentive with probability  $\kappa$ . The reminders then increase the probability of being attentive to  $\gamma + (1 - \gamma)\kappa$ . In the results below, I make the additional restriction that  $\kappa_t^1 = \kappa_{t+\Delta}^2$  for  $t \leq T_1$ , which allows me to examine how adding the same types and same number of cues to both the short and long deadline environments will affect completion rates and welfare.

**Proposition 9.** *Consider two deadlines  $T_1, T_2 = T_1 + \Delta > T_1$ . Then for any  $\{\gamma_t^{\kappa_t^i}(1, d)\}_{t \geq 1}$ ,  $\{\gamma_t^{\kappa_t^i}(0, d)\}_{t \geq 1}$  with the additional restriction that  $\kappa_t^1 = \kappa_{t+\Delta}^2$ :*

1. *Suppose that either Condition I( $T_1, T_2$ ) holds or that  $b_{T_2} + \underline{\xi} \geq 0$ . Then there is a  $\bar{\kappa} < 1$  such that  $\Pr(Q^{n, T_1} \leq T_1) < \Pr(Q^{n, T_2} \leq T_2)$  and  $\Pr(Q^{s, T_1} \leq T_1) < \Pr(Q^{s, T_2} \leq T_2)$  if  $\kappa_t^1 = \kappa_{t+\Delta}^2 \geq \bar{\kappa}$  for  $t \leq T_1$ .*
2. *Suppose that  $b_{T_2} + \underline{\xi} \geq 0$ . Then  $\Pr(Q^{n, T_1} \leq T_1) - \Pr(Q^{n, T_2} \leq T_2)$  is strictly decreasing in  $\kappa_{T_1}^1 = \kappa_{T_2}^2$ .*
3. *Suppose that Condition I( $T_1, T_2$ ) holds and the initial cue distributions are time invariant,  $\gamma_t^0 = \gamma^0$  for all  $t \geq 1$ . Then for each  $t \leq T_1$ ,  $\Pr(Q^{n, T_1} \leq T_1) - \Pr(Q^{n, T_2} \leq T_2)$  is strictly decreasing in  $\kappa_t^1 = \kappa_{t+\Delta}^2$  while it is positive.*

Proposition 9 motivates a simple and testable hypothesis for situations in which longer deadlines lead to lower completion rates, and in which payoffs from task completion don't change with time, as in the settings studied by Shafir and Tversky (1992), Shu and Gneezy (2010), Silk (2004), and Janakiraman and Ordóñez (2012). Suppose, for example, that study participants are more likely to complete a task when given a 2-day deadline than when given a 3-week deadline. Part 1 of Proposition 9 implies that if the participants receive a

sufficiently strong set of cues during the last two days in the 3-week condition and during both days of the 2-day condition, then the longer deadline should no longer lead to lower completion rates.

The second part of the proposition considers the case in which doing the task in the last period of the long deadline always generates positive flow utility that period. In this case, adding cues, however weak, in the last period of both the short and long deadline conditions should diminish the difference in completion rates between the short and long deadline conditions.

The last part of the proposition shows that under the additional assumption that the initial cue distributions do not vary from period to period, adding additional cues (however weak) any number of periods before the deadline has a bigger impact on a naive DM's completion probability when the deadline is long.

Attention decay is not only a sufficient condition, but also a necessary condition for reminder provision to have a larger effect in the long deadline condition. If the probability of being attentive to the task does not decline with time then, because more decision makers will have completed the task by period  $T_2 - 1$  than by period  $T_1 - 1$ , a reminder in period  $T_i - 1$  will actually have a *smaller* effect in the longer deadline condition. And even when the likelihood of being attentive does decay with time, the "dropout effect" described in the previous sentence is still a countervailing force that pushes cues to have a smaller net effect in the long deadline condition. The conditions described in Proposition 9 ensure that the "decay effect" dominates the "dropout effect."

Results such as Proposition 9 motivate a straightforward way to diagnose when certain patterns of behavior are caused by inattention: increasing attention through various types of cues should diminish those behavioral patterns. Although other biases such as procrastination may also play a role people's ability to complete tasks, tests such as those suggested by Proposition 9 can help assess *how much* of a certain behavioral pattern is caused by

inattention.<sup>38,39</sup>

## 1.4.2 Experimental Evidence

### Design and Procedures

Motivated by Proposition 9, I conducted an online, real-effort experiment to measure how much of the long deadline effect (previously documented in other studies) is due to the decay of attention over time. In addition to varying deadline length as in previous work, the new experiment also varied reminder provision to test the comparative static derived in Proposition 9. The experiment consisted of two phases described below.

Registration Phase (Day 0) Potential subjects received an invitation email to complete a 10-20 minute survey in which they would have to choose between hypothetical gambles, and for which they would receive a \$10 Amazon.com gift card. Those who were interested in the opportunity followed the link to the study site, where they created an account for the experiment. Subjects then answered a few demographics questions, were randomized into one of the four experimental conditions described below, and read instructions for completing the risk survey. Instructions (which included a link to the study site) were also automatically emailed to the subjects upon completion of the registration phase.

Task Completion Phase (Days 1–deadline) The risk survey became available to subjects at

---

<sup>38</sup>Short-run impatience (Laibson, 1997; O'Donoghue and Rabin, 1999) can lead to suboptimally low completion rates but, like the perfectly attentive and time consistent model, it does not predict that increasing the (finite) deadline can decrease completion rates when the distributions of completion payoffs do not change over time. The simple reason is that even if a longer deadline  $T_2$  induces more procrastination, upon reaching the last  $T_1$  periods of the longer deadline, the present-biased DM now plays the same game that he would play with the shorter deadline. Moreover, a fully attentive but present-biased DM would not exhibit suboptimally low completion rates for immediately pleasurable tasks as, arguably, in the experiments of Shu and Gneezy (2010).

There are, however, other deadline effects that are explained by present-biased preferences but not by the inattention model. Ariely and Wertenbroch (2002) study tasks that can only be completed over several days, such as term papers, and find that students self-impose shorter deadlines. This is consistent with students wanting to commit their short-run self to complete the task in a timely manner, but cannot be explained by inattention.

<sup>39</sup>Another theory that would be consistent with Proposition 7 but not Proposition 9 is a combination of declining motivation and projection bias (Loewenstein et al., 2003). For example, people may be excited when they first decide to complete some task, but their motivation to do it decays each day. If they misforecast this decay in motivation, then logic similar to that of Proposition 7 would predict that longer deadlines could lead to lower completion rates. However, reminders would then not reverse this trend, contrary to Proposition 9.

8:00 a.m. the day after registration. The four conditions for task completion were as follows:

1. *Short Deadline / No Reminders.* Subjects in this condition faced a deadline of 11:59 p.m. on day 2 and received no reminders. The study instructions clarified that they would receive no reminders, stating “You will receive no further communications from us over the course of the study.”
2. *Long Deadline / No Reminders.* Subjects in this condition faced a deadline of 11:59 p.m. on day 21. As in the previous condition, they received no reminders on any of the days, and the study instructions clarified that they would not.
3. *Short Deadline / Reminders.* Subjects in this condition faced a deadline of 11:59 p.m. on day 2, and received reminders on days 1 and 2. Subjects were informed that they would be getting reminders each day. The reminders are described below in further detail.
4. *Long Deadline / Reminders.* Subjects in this condition faced a deadline of 11:59 p.m. on day 21, and received reminders on days 20 and 21. Subjects were informed that they would be getting reminders on days 20 and 21.

The timing and content of the reminders were as follows. Reminders were only sent to those subjects in the reminders conditions who had not yet completed the survey. Of those subjects, email reminders were sent to everyone. Additionally, during the registration phase, subjects randomized into the reminders conditions were also given the option to consent to receive text message reminders. All subjects in the reminders conditions who had not yet completed the survey received an email reminder at 8 a.m. on each of the two reminder days. Subjects were sent a second set of reminders at either 10 a.m., 11 a.m., 12 p.m., 1 p.m., or 2 p.m. that consisted of an email and, for those choosing to receive text messages, a text message. Two-thirds of all subjects in the reminders condition received only *basic* email reminders. These emails mentioned the deadline date and nothing else (see Appendix A.3 for text). One-third of the subjects received a *basic+info* email reminder instead of a basic

email reminder at 8 a.m. of each reminder day. The purpose of the augmented message was to examine whether information loss played a role in task completion. The email therefore included information about the \$10 reward and when it would be emailed, the length of the survey, and the URL of the study (see Appendix A.3). The language of the reminder messages was intentionally simple to minimize demand effects.

After completing the risk survey, subjects were prompted to sign a receipt for the \$10 gift card, which was always emailed the day after the deadline.<sup>40</sup>

## Results

A total of 403 subjects (93% students; 88.6% Harvard undergraduates) were recruited for the experiment in August and September. The email was sent to all 12 undergraduate residences at Harvard<sup>41</sup> and generated a total of 34 unique dates on which subjects signed up for the study. Of subjects in the reminders condition, 64.5% chose to receive text messages. Appendix A.3.2 verifies that the randomization was successful, finding no significant differences between any of the demographic variables across conditions.

Table 1.2 summarizes the completion rates by condition; and table 1.3 estimates a linear probability model of task completion, with robust standard errors clustered by both start date and undergraduate residence. Consistent with previous evidence and the prediction of Proposition 7, a longer deadline decreases completion rates from 59.4% to 41.6% in the no reminders conditions ( $p < 0.01$ ). Consistent with Proposition 9, however, reminders close

---

<sup>40</sup>Note that the reward was never immediately available. Nevertheless, one potential concern with this aspect of the procedures is that if subjects have a high exponential daily discount factor, then the discounted value of the reward for subjects on day 1 in the long deadline conditions is lower than the discounted value of the reward for subjects on day 1 in the short deadline conditions. However, all results about completion rates would continue to hold even with exponential discounting—the simple intuition is that the discounted value of task completion on days 20 and 21 in the long deadline conditions is equal to the discounted value of task completion on days 1 and 2 in the short deadline conditions. At the same time, to the extent that small amounts of money are fungible, subjects should not treat the \$10 on day 22 any differently than the \$10 on day 3. For small stakes monetary rewards, Andreoni and Sprenger (2012) experimentally estimate an annual discount rate of 0.3, which translates into a 3-week discount rate of less than 1.7%, or a discount of 0.17 from \$10.00.

<sup>41</sup>Almost all sophomores, juniors and seniors at Harvard live in one of the 12 undergraduate “houses.” A recruitment email was sent to each house only once. Recruitment occurred over the course of 5 weeks, with 2 or 3 houses receiving a recruitment email each week.

the gap between the short and long deadlines. With reminders, the longer deadline reduces completion rates by an insignificant 2.3%.

**Table 1.2:** *Fraction of Subjects Completing Task, by Experimental Condition*

	No reminders	Reminders on last two days
2 days	59.4%	74.3%
3 weeks	41.6%	72.0%

**Table 1.3:** *Probability of Completing Task, by Experimental Condition*

Pr(Complete)	(1)
LongDeadline	-0.178*** (0.067)
Reminders	0.149** (0.072)
LongDeadline*Reminders	0.156* (0.081)
Adj. $R^2$	0.074
Observations	403

*Notes:* This table estimates a linear probability of task completion, by condition. The variable “LongDeadline” equals 1 if subjects had 3 weeks to complete the task, and equals 0 if subjects had 2 days to complete the task. The variable “Reminders” equals 1 if subjects received two days of reminders—days 1 and 2 for subjects with a 2-day deadline and days 20 and 21 for subjects with a 3-week deadline. Robust standard errors are computed by specifying both start date and undergraduate residence as the cluster groups, following the multiway clustering method suggested by Cameron et al. (2011). \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

The key prediction that is confirmed by the data is that *if* longer deadlines decrease completion rates, *then*, because of how rapidly the likelihood of being attentive must decay, reminders have a larger impact on decision makers facing a longer deadline. Table 1.3 shows that while subjects given the *short* deadline are 14.9 percentage points more likely to complete the task when they receive reminders, subjects facing the *long* deadline are 30.5 percentage points more likely to complete the task when they receive reminders. The difference between differences is a marginally significant 15.6 percentage points ( $p < 0.1$ ).

Table A.7 in Appendix A.3 shows that these results are robust to both demographic

controls as well as day of week controls.<sup>42</sup> Table A.8 in Appendix A.3 analyzes whether different types of reminders had different effects. The content of the reminder had no additional effect on completion rates, suggesting that information loss is not an important factor. The table also shows that the 64.5% of the subjects who agreed to receive SMS messages did not behave any differently from the subjects who did not. This fact, of course, is difficult to interpret because of the usual endogeneity caveats.

Figure 1.4 provides further evidence for the decaying attention dynamics proposed by the model. For each of the four experimental conditions, the figure shows the fraction of subjects completing the task on a given day. Consistent with the theoretical prediction that longer deadlines increase the perceived option value of delaying task completion, subjects in the long deadlines conditions are significantly less likely to complete the task on the first day that it becomes available. Consistent with decaying attention, however, subjects who don't complete the task in the first few days are very unlikely to ever complete it if they don't receive reminders. Quite strikingly, only 2 out of the 101 subjects in the *long deadline / no reminders* condition completed the task between days 11 and 21. Subjects in the *long deadline / reminders* condition exhibited a similar pattern of behavior during days 1-19. When they received reminders on days 20 and 21, however, more than 50% of the subjects who had not yet completed the task ended up completing it.<sup>43</sup>

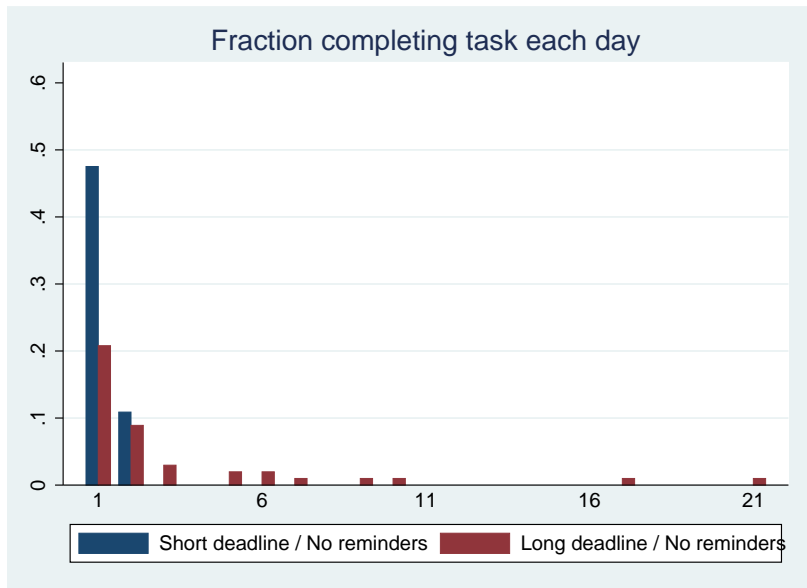
The results of this experiment thus replicate existing studies on deadline effects but, by showing how a small set of reminders can drastically change the outcomes, provide strong evidence that these deadline effects are generated by the decaying attention mechanism postulated in this paper.

---

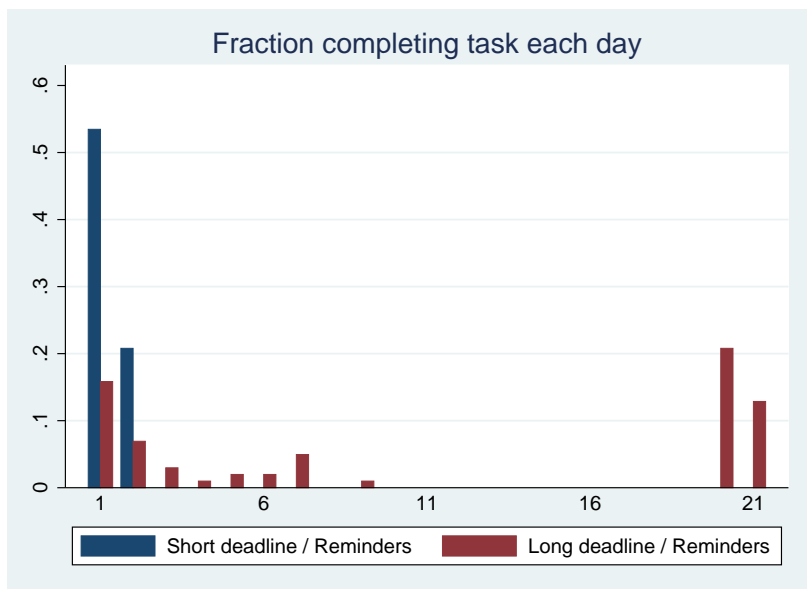
<sup>42</sup>Recruitment was staggered across different days of the week, so that each day of the week served as a start date for at least 33 and no more than 77 subjects in the study. Table A.6 shows that deadline day of week appears to have no impact on probability of completing the task: the hypothesis that there are no day of week effects cannot be rejected for either the short deadline or the long deadline conditions ( $p = 0.18$  for short deadline;  $p = 0.26$ ) for long deadline).

<sup>43</sup>This striking pattern of behavior also shows that reminder provision is not likely to have generated a demand effect. One potential hypothesis is that reminders induce a demand effect by signaling to subjects that the experimenter is taking the task seriously. Note, however, that subjects knew from the beginning whether or not they were getting reminders. The extremely low completion rates in days 11-19 in the long deadline/reminders condition are not consistent with this signaling demand story.





(a) No Reminders



(b) Reminders

Figure 1.4: Fraction of Subjects Completing Task on Each Day, by Condition

## 1.5 Inattention in the Market

In this section, I explore how the kinds of behavioral patterns analyzed in Sections 1.3 and 1.4 might play out in various strategic interactions between sophisticated firms and inattentive consumers. Section 1.5.1 builds on the task completion results from Section 1.4 to study consumer rebates and related applications, while Section 1.5.2 builds on the repeated action results from Section 1.3 to study a model of reminder advertising.

### 1.5.1 Rebates and Related Applications

In this section I illustrate the applicability of the inattention model by embedding it in a simple model of consumer rebates. Rebates are a commonly used marketing tactic that requires consumers to 1) make a purchase at some up-front price  $p$  and 2) to submit a (partial) refund request, along with proof of purchase, by mail or internet by some deadline  $T$ . Edwards (2007) estimates the annual volume of rebate offers to range from \$4 to \$10 billion. Silk and Janiszewski (2008) estimate redemption rates of 10% to 30% for consumer rebate offers between \$20 and \$100, and redemption rates of 40% for bigger ticket consumer electronics.

Previous work on consumer rebates has focused on their potential use for price discrimination (Narasimhan, 1984; Banks and Moorthy, 1999; Lu and Moorthy, 2007),<sup>44</sup> and has modeled consumers' redemption decision as a static, one-period choice. The analysis here differs from previous work in two crucial ways. First, the analysis is the first, to my knowledge, to investigate the question of how firms choose optimal redemption deadlines in a dynamic model of consumers' redemption decisions. Second, I investigate a setting in which firms may offer rebates to naive and inattentive consumers even when there is no incentive to price discriminate among them. While price discrimination is surely a potential motivation as well, the intentionally simple setting in this section allows me to formally

---

<sup>44</sup>Gerstner and Hess (1991) propose that rebates may arise due to channel distribution issues between manufacturers, retailers and consumers. Chen et al. (2005) propose that rebates may function as "state-contingent discounts" that redistribute money to states with high marginal utility from income. More similar to the analysis here is Gilpatric's (2009) recent work on offering rebates to time-inconsistent consumers.

capture and evaluate policymakers' and industry leaders' claims that consumer rebates are deceptively attractive to consumers.<sup>45</sup>

The monopolistic seller produces a certain product at a constant marginal cost  $c$ , and in period 0 chooses an offer  $\mathcal{P} = (p, T, r)$  consisting of an upfront price  $p$ , a rebate  $r \geq 0$  and a redemption deadline  $T$  chosen from a set of possible deadlines  $\mathcal{T}$ . For simplicity, I assume that  $\mathcal{T}$  is finite. The set  $\mathcal{T}$  may be bounded from below because of a minimum deadline length requirement as in, e.g., New York state.<sup>46</sup>, and bounded from above because of similar restrictions<sup>47</sup> or because of logistical and credibility issues.

I will often focus on the monopolist's optimal offer conditional on being restricted to a particular deadline  $T$ .<sup>48</sup> This is a helpful intermediate step in the analysis that also facilitates the evaluation of various policy proposals to restrict deadline choice. The optimal policy for a particular deadline  $T$  is denoted  $\mathcal{P}_T^* = (p_T^*, r_T^*)$ , and the corresponding monopolist's profits are denoted  $\pi_T^*$ . The optimal policy  $\mathcal{P}$  is computed by choosing the deadline  $T \in \mathcal{T}$  that maximizes  $\pi_T^*$  and then setting  $p$  and  $r$  according to  $\mathcal{P}_T^*$ .

There is a unit mass of consumers who are interested in the product and value it at  $v > 0$ . They make a purchase decision in period 0. In each period  $t = 1, \dots, T$ , consumers who have purchased the product can incur an i.i.d. hassle cost  $\xi_t \sim F$  (with  $\xi_t \leq 0$ ) and mail in the rebate. Consumers who never attempt to redeem the rebate derive utility  $v - p$  from the transaction. Consumers who mail in the rebate in period  $t$  derive expected utility  $v - p + \theta r + \xi_t$  from the transaction, where  $\theta < 1$ . The discount factor  $\theta$  arises from the possibility of filling out the form incorrectly and thus not receiving the rebate (Edwards, 2007, 2009) or the possibility of losing the check once it arrives in the mail (Edwards, 2007,

---

<sup>45</sup>Preliminary results show that even when price discrimination is possible, making rebates maximally deceptive is more profitable than trying to price discriminate among consumers.

<sup>46</sup>See Edwards (2009) for a summary of the law.

<sup>47</sup>To my knowledge, no such restrictions exist. However, the theoretical analysis in this section will consider the possibility of setting an upper bound.

<sup>48</sup>Proposition 10 shows that an optimal offer always exists. An optimum could fail to exist if it was the case, for example, that the monopolist's profits increase without bound for an appropriate combination of  $p$  and  $r$  approaching infinity.

2009).<sup>49</sup> Accordingly, the monopolist's profit is given by  $\pi = p - c - \theta\mu r$ , where  $\mu$  is the probability that the consumer attempts to redeem over the course of the  $T$  periods.

Following the recent behavioral industrial organization literature on two-part pricing (Heidhues et al., 2012b,a; Grubb, 2012) I assume a price floor: there is some  $\underline{p}$  such that any offer must satisfy  $p - r \geq \underline{p}$ . Following (Heidhues et al., 2012b, 2011), this can be motivated by the existence of market-savvy arbitrageurs who have no interest in the actual product, but who will exploit a firm offering easy money. Appendix A.1.5 walks-through this microfoundation.<sup>50</sup> For smaller value items,  $\underline{p} = 0$  is consistent with, and helps explain the fact that "free after rebate" deals are extremely common, while deals in which rebates are even slightly above up-front prices are non-existent.<sup>51</sup> All results will be proven for any  $\underline{p} < c$ , however.<sup>52</sup>

I begin by characterizing when rebates will be offered.<sup>53</sup>

**Proposition 10.** *For each  $T$ , there exists an optimal policy  $\mathcal{P}_T^*$ , and consumers always buy in equilibrium if  $v \geq c$ . If consumers are sophisticated about their inattention then  $p_T^* = v$  and  $r_T^* = 0$ . If consumers are naive, then for any set of attention probabilities  $\gamma_t(\alpha, x)$ , there is a  $v^\dagger > 0$  such that*

1. *If  $v < v^\dagger$  then  $r_T^* = 0$  in any equilibrium in which consumers buy the product.*
2. *If  $v > v^\dagger$  then  $p_T^* > v$  and  $r_T^* > 0$  and consumers buy the product.*

---

<sup>49</sup>The results in this section would be unchanged if instead I assumed that 1) there is some probability that consumers lose the rebate form or the necessary information they need to claim the rebate (e.g., barcodes) or if 2) consumers are partially naive but not fully naive. What matters for the analysis in this section is that all consumers, including the naive ones, discount the value of the rebate because of a mistake they may make.

<sup>50</sup>And as I also discuss in the appendix, all results hold under the plausible assumption a small fraction of the consumers who derive value from the product happen to be market-savvy arbitrageurs. As Heidhues et al. (2011) point out, these arbitrageurs are equivalent to Gabaix and Laibson's (2006) *sophisticates*.

<sup>51</sup>See, for example, <http://www.dealigg.com/free.php>. Of course, current rebate offers still don't keep out all arbitrageurs. Woodruff (2012), for example, describes profit-making schemes involving the resale of free or nearly free after rebate offers.

<sup>52</sup> $\underline{p} = 0$  may not be a good assumption for items such as expensive consumer electronics that can be resold in a secondary market. This suggests that a more general model could allow the price floor to be a function of  $v$ . All results in hold for a piece-wise linear formulation of the form  $\underline{p}(v) = \max(0, a + bv)$ , where  $b \in (0, 1)$ .

<sup>53</sup>As before, I assume that there is some  $\bar{z} < 1$  such that  $\gamma_t(1, d) \leq \bar{z}$  for all  $t$ .

According to Proposition 10, rebates will not be offered to sophisticated consumers. The simple intuition is that the monopolistic seller extracts all surplus generated by a transaction with a sophisticated consumer. Offering a rebate, however, reduces total surplus because it forces consumers to incur effort costs to redeem it.

With naive consumers, the monopolist's profits are no longer determined solely by the transaction surplus; the profits are given by the sum of the actual transaction surplus *and* consumers' misperceptions of the value of the deal. Thus a rebate will be optimal when it increases misperceptions more than it decreases the actual surplus. Note now that the introduction of a rebate cannot decrease true surplus by more than  $E_F \zeta_t$ , the expected redemption effort cost in a given period. In contrast, holding  $T$  fixed, the misperceptions of naive consumers will strictly increase in  $r$ . But as shown in Appendix A.4.3, the post-rebate price floor constrains the rebate value to be no higher than some continuous function  $\bar{n}(v)$  that is strictly increasing in  $v$  and satisfies  $\bar{n}(0) = 0$ . Thus a high  $v$  is necessary (and sufficient) to guarantee that the monopolist can set a rebate that will generate sufficiently high misperceptions.

Because rebates are used to create deceptively appealing offers, expected consumer utility is negative in any equilibrium with a rebate. The next proposition shows that rebates may also lead to socially wasteful transactions in the sense that a product that costs  $c$  to produce is sold to consumers who derive utility  $v < c$  from the product.

**Proposition 11.** *Suppose consumers are naive. Then for each  $L > 0$  and  $T$  there exists  $c^\dagger$  such that if  $v = c - L$  and  $c \geq c^\dagger$  then  $p_T^*, r_T^* > 0$  and consumers purchase the product.*

The intuition behind proposition 11 is again derived from the fact that the monopolist's profits are given by the sum of the actual transaction surplus and consumers' misperceptions of the value of the deal. Thus even when  $v < c$  and the transaction surplus is negative, a rebate can still be profitable by making the deal appear deceptively attractive to consumers. As with Proposition 10, the price-floor requires the product to be a sufficiently big ticket item to create scope for sufficiently deceptive rebates.

I now characterize how the monopolist will choose a redemption deadline and how that

will impact consumers.

**Proposition 12.** *Let  $R(r, T)$  denote the probability that a consumer mails in a rebate of size  $r$  given a deadline  $T$  and suppose that consumers are naive.*

1. *If  $T_1 < T_2$  and  $R(r_{T_1}^*, T_1) < R(r_{T_1}^*, T_2)$  then  $\pi_{T_2}^* > \pi_{T_1}^*$ .*
2. *Assume there is some  $\underline{z} > 0$  such that  $\gamma_t(0, d) \geq \underline{z}$  for all  $t$ . Then holding all other parameters constant, there is a  $T^\dagger$  such that  $r_T^* = 0$  for all  $T \geq T^\dagger$ .*

The main message of Proposition 12 is that if consumers are sufficiently inattentive, firms will choose deadlines that are of “intermediate length.” This basic message follows from Proposition 7, which shows that longer deadlines can lead to lower task completion rates, and from Proposition 8, which shows that long enough deadlines will actually lead to higher completion rates than “intermediate length” deadlines. Although a systematic analysis of redemption deadlines is lacking, the “medium length” deadline prediction seems consistent with industry practice of setting deadlines that typically range between 15 and 60 days.<sup>54,55</sup>

The model’s prediction that rebates may be used to create deceptively attractive deals is echoed by industry experts and policymakers. As bluntly explained by a VP of an electronics retailer, “Manufacturers love rebates because the redemption rates are close to none. It’s just human nature that we go after them, and they get people into stores, but when it comes time to collect, few people follow through.”<sup>56</sup> This prediction is in contrast to the more traditional views that rebates are used to price discriminate among rational consumers. And even more starkly in contrast to the rational price discrimination framework is the prediction that rebates can facilitate the sale of socially wasteful products.

---

<sup>54</sup>See, however, the caveats in footnote 37

<sup>55</sup>Rebates for products offered by the retailer Staples, for example, typically have redemption deadlines of either 30 or 60 days (see [https://www.stapleseasyrebates.com/promocenter/staples/promo\\_search.html](https://www.stapleseasyrebates.com/promocenter/staples/promo_search.html)). Silk and Janiszewski (2008) analyze a random online sample and find that most deadlines are either 15 or 30 days.

<sup>56</sup>Catherine Greenman, “The Trouble With Rebates.” The New York Times. September 16, 1999. Accessed August 15, 2013, <http://www.nytimes.com/1999/09/16/technology/the-trouble-with-rebates.html?pagewanted=all&src=pm>

While some have explored the possibility of banning rebates on the grounds that they are deceptive (Lynch and Zauberger, 2006; Sovern, 2006), the model suggests that there are less paternalistic interventions that can increase consumer welfare or market efficiency, without interfering with the potential of using rebates for purposes of efficiency-enhancing price discrimination. One such possibility, for example, is to require very short or very long deadlines. Legislators in the states of California, New York, Texas, and North Carolina have imposed minimum deadline lengths (Edwards, 2007), though it is unclear whether those 2 or 4 week minimums are sufficiently high. Note, however, that such deadline restrictions have less clear-cut effects on social efficiency. This is because by virtue of attempting to minimize redemption probability, a monopolist's choice of deadline also minimizes consumers' socially wasteful redemption effort.<sup>57</sup>

Lastly, note that the inattention model also suggests that because of attention decay, exploitative firms may want to create *deferred rebate programs* that require consumers to wait before they can claim the rebate. Some companies have, in fact, offered 54-week deferred rebates of up to \$10,000 to consumers purchasing big ticket items such as pools or automobiles, but have subsequently faced lawsuits for deceptive and exploitative practices.<sup>58</sup> Following incidents of massive consumer complaints, North Carolina has banned rebates with a deferral period of more than 6 months, in addition to creating other restrictions.<sup>59</sup> I speculate that deferred rebates are not more common because firms fear that such offers are so transparently exploitative that they would tarnish the firms' reputations on the one hand, and—as in the case of North Carolina—invite more severe government oversight and intervention on the other hand.

---

<sup>57</sup>A second possibility is to minimize inattention bias by requiring firms to send consumers reminders to redeem their rebates. A third possibility is to attempt to debias naive consumers through salient disclosure of low redemption rates (Edwards, 2007), or by encouraging them to set their own reminders to minimize the probability of forgetting.

<sup>58</sup>Jamie Boll, "PSI: Big rebate offer leads to big disappointment." Feb 08, 2011. Accessed 8/25/2013, <http://www.wbtv.com/story/13983724/psi-big-rebate-offer-leads-to-big-disappointment>. See also the Pennsylvania Attorney General Press release: <http://www.attorneygeneral.gov/press.aspx?id=835>

<sup>59</sup>See the North Carolina Department of Justice on rebates: <http://www.ncdoj.gov/Consumer/Purchases-and-Contracts/Rebates.aspx>

## Related Applications

There are a number of other transactions between firms and consumers that share some key similarities with consumer rebates. Namely, the feature that, at a later date, the consumer can take some action  $x$  by a deadline  $T$ , which generates some (endogenously set) benefit  $b - \zeta_t$  to the consumer, at a (endogenously set) cost  $c$  to the firm. A product return policy with a deadline  $T$  is one such example. Another set of examples is contracts with automatic renewal. DellaVigna and Malmendier (2004) list examples such as free trial offers that automatically transition into paid services, automatically renewing contracts in the health club industry, and automatically renewing newspaper subscriptions. In all of these examples, naive and inattentive consumers will overvalue the deal because they will overestimate the likelihood of returning the product or canceling their membership.

As DellaVigna and Malmendier (2004) show, naive hyperbolic discounting can also lead to overvaluation, and can explain why firms increase switching costs by, for example, stipulating that consumers can only cancel their gym membership in person. The analysis in this paper provides an alternative explanation: by increasing switching costs, firms increase the option value of delay, which makes naive and inattentive consumers more likely to put off the task until later and subsequently lose track of it. Intuitively, if canceling a gym membership involved only one click online, a consumer who first decided to cancel the membership would likely do it right away. The requirement to cancel the membership in person, however, makes the consumer put off the task until later and subsequently forget about it.

The inattention model can thus explain why firms make some actions very easy—like the “1-Click” buying feature—while making other actions more effortful. But additionally, the inattention model also makes predictions about which actions firms will remind consumers to take, and which they won’t. Firms exploiting naive inattention would remind consumers to, for example, follow through on limited time offers, but would not remind consumers to, for example, consider canceling their subscription after the expiration of the free-trial phase. These predictions seem to be consistent with casual observation, and future work should



provide systematic evidence for these new predictions.

### 1.5.2 Optimal Cue Provision by an Interested Party

Economic theories of advertising typically assume that marketing communications either provide information (Stigler, 1961; Butters, 1977) or enter directly into consumers' utility functions (Becker and Murphy, 1993). Both the informative and the persuasive (or complementary) functions of advertising are important, particularly when the product is first introduced. However, marketers also emphasize a third function of marketing communications: after a product has been introduced, advertising must keep the product at the top of the consumer's mind.

This section uses the inattention model to formalize a model of reminder communications by a sophisticated firm or organization. To keep things simple, I focus on an organization whose objective is to direct consumer behavior *solely* through the reminder communications. For example, a health care provider or insurer might use SMS messages or phone calls to remind patients with chronic diseases to take their medications,<sup>60</sup> or an organization might send reports reminding consumers of ways to save energy (Allcott and Rogers, 2012). The insights developed in this simple model could also be applied to understand the optimal policy of a firm that uses reminder advertising to make sure that its repeat purchase product stays a top of mind consideration.

While recent work in economics has investigated advertising to inattentive consumers in a static, one-period setting (e.g., Falkinger 2007, 2008; Eliaz and Spiegler 2011a,b),<sup>61</sup> the inattention model proposed in this paper provides foundations for extending these analyses to dynamic, multi-period environments. As this section shows, some of the most distinctive predictions of reminder communications are manifested only in dynamic settings.

---

<sup>60</sup>Providers and insurers may purchase these services from independent third party firms specializing in population health and patient communication; e.g., Healthways, Phytel, Optum, IncentOne, Staywell.

<sup>61</sup>See also Hefti (2013) for a model building on Falkinger (2007, 2008). See Spiegler (forthcoming) for a model of competitive framing. Most related to the analysis here, see Villas-Boas (1993) for a simple model of reminder advertising in a duopoly.

Formally, there is a unit mass of homogeneous, inattentive consumers who choose  $x_t \in \{d, a\}$  in each period  $t \geq 1$ . Each period, the organization chooses between sending a communication, denoted  $m_t = 1$ , or not sending a communication, denoted  $m_t = 0$ . A period  $t$  communication costs  $c$  to send and reaches a consumer with probability  $w \in (0, 1)$ . For simplicity, I assume that when a consumer is reached he is attentive with probability 1 regardless of previous history. Otherwise, there are no other cues and the DM is attentive with probability  $g(\alpha, x, 0)$ . For simplicity, I assume that  $g(0, d, 0) = 0$ , and that the probability of being attentive in period 0 is 0.

The organization's objective function is to choose  $m = (m_1, m_2, \dots)$  to maximize  $E_0 \sum_{t=1}^T \delta_o^t (\phi_t - m_t c)$ , where  $\phi_t$  is the fraction of consumers choosing  $x_t = a$  and  $\delta_o < 1$  is the organization's discount factor. Formally, I will assume that the organization commits to a strategy in period 0. However, the assumptions in this section will guarantee that the consumers will not alter their strategies in anticipation of future cues; thus, the results would be identical under the assumption that the organization can revise its strategy each period.

Note that initially uninformed but otherwise attentive consumers, in the sense of Stigler (1961) or Butters (1977), are nested as a special case of the framework:  $g_0(0, d, 0) = 0$  and  $g(1, d, 0) = 1$ . That is, these consumers will not consider  $x_t = a$  until they are contacted by the organization, but will always consider  $x_t = a$  after the first time. The general framework proposed here thus nests a simple variation of informative advertising as a special case.

To fully draw out the long-run implications for the optimal messaging strategy, I extend the baseline model and set  $T = \infty$ , and assume that future payoffs are discounted by an exponential discount factor  $\delta_{DM}$  (see Appendix A.1.1 for a formal extension of the inattention model to infinite horizons). Each consumer's period  $t$  payoff from choosing  $x_t = a$  is given by  $b_H > 0$  with probability  $\ell \in (0, 1)$  and by  $b_L < 0$  with probability  $1 - \ell$ . These taste variations are independently distributed across the unit mass of consumers each period.<sup>62</sup>

---

<sup>62</sup>To map this into the formal framework, assume that  $b_t \equiv b$  for all  $t$ , and suppose that  $\zeta_t$  takes on the values  $\zeta_H$  or  $\zeta_L < \zeta_H$ . So  $b_H = b + \zeta_H$  while  $b_L = b + \zeta_L$ . Note that in contrast to the assumptions in earlier sections, the distribution  $F$  is not atomless here.

To further simplify the model, I assume that  $b_L$  is sufficiently low so that even sophisticated consumers always choose  $x_t = d$  when  $b_L$  is realized.<sup>63</sup>

Proposition 13 establishes necessary and sufficient conditions for repeated communication to be optimal in the long run.

**Proposition 13.** *Let  $m^*$  be an optimal messaging strategy and set  $\psi = \ell g(1, a, 0) + (1 - \ell)g(1, d, 0)$ .*

1. *If*

$$\frac{w\ell}{1 - \delta_o\psi} \leq c$$

*then  $m_t^* = 0$  for all  $t$ .*

2. *If  $\psi = 1$  then there exists a  $t^\dagger < \infty$  such that  $m_t^* = 1$  if and only if  $t < t^\dagger$ .*

3. *If*

$$\frac{w\ell}{1 - \delta_o\psi} > c \tag{1.10}$$

*and  $g(1, a, 0) < 1$ , then for any  $t > 0$  there exists a  $t' \geq t$  such that  $m_{t'}^* = 1$ .*

Part 1 establishes conditions under which sending communications to the consumers is not optimal in any period. The quantity  $\frac{w\ell}{1 - \delta_o\psi}$  is an upper bound on how much a single message can increase the organization's discounted payoff, so that sending a message is never optimal when this upper bound is not greater than  $c$ . As intuition would suggest, the upper bound is increasing in  $\ell$ , the consumers' likelihood of wanting to choose  $x_t = a$  conditional on being attentive, and increasing in  $w$ , the efficacy of the organization's messaging attempt.

Note that in the absence of rehearsal effects,  $\psi = 0$ , and the upper bound reduces to  $w\ell$ . With rehearsal effects,  $\psi > 0$ , and thus a single period  $t$  message has a larger effect:

---

<sup>63</sup>A sufficient condition for this is that  $b_L < -\gamma b_H / (1 - \psi)$ , where  $\psi = \ell g(1, a, 0) + (1 - \ell)g(1, d, 0)$ . Note that while the parameter  $\ell$  is a useful comparative statics parameter corresponding to the consumers' strength of preference for choosing  $x_t = a$ , the variation in taste shocks does create additional complexity arising from the fact that sophisticated consumers' choice may be shaped by their beliefs about the organization's future messaging strategy.

in addition to making DMs more attentive in period  $t$ , the message also makes DMs more attentive in periods  $t + 1, t + 2, \dots$ .

Part 2 establishes conditions under which sending communications to the consumers cannot be optimal in the “long-run,” though may be optimal in the short run. The condition that  $\psi = 1$  simply ensures that the probability of being attentive does not decay with time. Importantly, this condition holds for consumers with  $g_0(0, d, 0) = 0$  and  $g(1, d, 0) = 1$ —which can be interpreted as initially uninformed but otherwise attentive consumers.

Part 3 completes the proposition by showing that when neither the condition in part 1 nor the condition in part 2 holds, the organization’s optimal strategy will involve messaging in the long run.

Proposition 13 illustrates a key difference between optimal communications with inattentive versus attentive but uninformed consumers: repeated communication is optimal in the long run if and only if consumers are inattentive *and* their attention decays with time in the absence of communications. In fact, parts 1 and 3 imply that when  $\psi < 1$ , so that attention decays over time, the organization’s optimal advertising strategy is to either not send messages at all, or to never stop sending messages.<sup>64</sup>

But while the inattention model implies that the returns to additional communication do not converge to zero in the long run, like models of informative advertising it also predicts that the returns to each additional communication are diminishing. A version of this intuition was already established in part 5 of Proposition 2, which showed that temporally separated cues are substitutes. In the simple setting here, it is similarly true that for  $t' < t$ , the returns to choosing  $m_t = 1$  are decreasing in  $m_{t'}$  because a choice of  $m_{t'} = 1$  decreases the fraction of consumers who would otherwise be inattentive in period  $t$ . Such intermittent communications create an “action and backsliding effect”: for  $t > t'$ , the

---

<sup>64</sup>Note that models of informative advertising can sustain long-run advertising policies in a market by assuming overlapping generations of consumers, but these models also predict that an organization or firm will eventually want to limit its communications to zero with any one particular household or consumer. This distinction is particularly important for communications such as those encouraging health-improving or energy conservation behaviors, which are targeted to specific households and can be sent on a frequent and regular basis.

probability of  $x_{t'} = a$  is high following  $m_{t'} = 1$ , then decays while  $m_{t'+1}, m_{t'+2}, \dots = 0$ , and then increases substantially again when  $m_t = 1$ .<sup>65</sup> Both the diminishing returns prediction and the action and backsliding prediction are key determinants of the optimal advertising policy and both have received empirical support. Allcott and Rogers (2012), for example, find that the effect of each additional report about energy-saving behaviors declines with time. At the same time, they also find a significant action and backsliding effect.

Lemma 1 formalizes the conditions under which choosing  $m_t = 1$  in each period is not optimal. For the remainder of the section, I will use  $\mathbf{1} = (1, 1, \dots)$  to denote the strategy of sending a message each period, and likewise use  $\mathbf{0}$  to denote the strategy of never sending a message.

**Lemma 1.**  *$m = \mathbf{1}$  is an optimal messaging strategy if and only if*

$$\frac{w(1-\psi)\ell}{(1-\psi(1-w))(1-\delta_o\psi(1-w))} \geq c. \quad (1.11)$$

Combined, Proposition 13 and Lemma 1 begin to provide insights into how consumers' preferences for choosing  $x_t = a$ , parametrized here by  $\ell$ , affect the organization's returns to choosing  $m_t = 1$ . When consumers are attentive but initially uninformed, the returns to sending an additional message are increasing in  $\ell$ , for the simple reason that conditional on considering  $x_t = a$ , more consumers will choose it. Similarly, when consumers are inattentive but rehearsal has no effect,  $\psi = \ell g(1, a, 0) + (1 - \ell)g(1, d, 0) = 0$  and thus condition (1.10) and condition (1.11) are equivalent. This then implies that there is some  $\ell^\dagger$  such that the optimal strategy is  $m = \mathbf{1}$  for  $\ell > \ell^\dagger$  and  $m = \mathbf{0}$  for  $\ell < \ell^\dagger$ . Lemma 2 summarizes these results:

**Lemma 2.** 1. *If  $g(1, d, 0) = 1$  then the total number of messages is increasing in  $\ell$ .*

2. *If  $g(1, a, 0) = 0$  then the optimal strategy is  $m = \mathbf{0}$  if  $w\ell < c$  and is  $m = \mathbf{1}$  if  $w\ell > c$ .*

Behavioral rehearsal adds an additional nuance to the relationship between preferences

---

<sup>65</sup>Agarwal et al. (2013) first coined the phrase "action and backsliding" to describe this kind of pattern of behavior with respect to credit card fees. Their finding that behavior responds most strongly to cues in the form of late fees, however, is not directly explained by the model in this paper.

and the optimal messaging strategy. On the one hand, the higher the  $\ell$ , the more likely consumers are to choose  $x_t = a$  conditional on being attentive. On the other hand, higher  $\ell$  also implies that consumers are more likely to be attentive due to rehearsal. The combination of these two effects implies that the optimal communication intensity can be non-monotonic in  $\ell$ :

**Proposition 14.** *Fix  $g(1, d, 0) < 1$  and  $c < 1/4 - g(1, d, 0)/4$ . Then there exist high enough  $g(1, a, 0) < 1$  and  $w < 1$  such that the optimal messaging strategy  $m^* = (m_1^*, m_2^*, \dots)$  is non-monotonic in  $\ell$ : for some  $\ell_1 < \ell_2 < \ell_3$ ,*

*i) If  $\ell \in (\ell_1, \ell_2)$  then condition (1.10) holds but condition (1.11) does not hold*

*ii) If  $\ell \in (\ell_2, \ell_3)$  then condition (1.11) holds*

*iii) If  $\ell \in (\ell_3, 1)$  then condition (1.10) holds but condition (1.11) does not hold*

The key implication of Proposition 14 is that when the effect of behavioral rehearsal is sufficiently strong, cue provision has the biggest impact on consumers who want to choose  $x_t = a$  “often enough” but not necessarily all of the time. Intuitively, consumers who want to choose  $x_t = a$  all of the time develop a strong enough habit that will not be significantly affected by cues. In the daily action experiment, for example, participants who completed the survey for three weeks straight were not significantly affected by week 3 cues. On the other hand, consumers who rarely want to choose  $x_t = a$  will also not be affected by additional cues for the simple reason that they will not want to choose  $a$  conditional on being attentive. These results generate implications not only for how the organization should design its optimal messaging strategy for a given population of consumers, but also for how the organization should attempt to target its messages to different subgroups within a population.<sup>66</sup>

---

<sup>66</sup>The model also generates insights about how to optimally tailor communications to a given consumer over time. The results in Section 1.3 show that all else equal, reminders should be most effective following an interruption to the decision maker’s routine. An organization that can effectively monitor each consumer’s behavior should thus increase the intensity of communications following what appear to be random breaks in the repeated behavior. Such targeting should be especially affective when consumers are naive, because naive consumers do not change their behavior in anticipation of future cues.

Although the analysis in this section focuses on a simple setting in which the organization only chooses cues, an important question for future theoretical and empirical work is how to optimally combine incentives and cues. The substitutes and complements predictions in Propositions 3 and 4 suggest that when both incentives and cues part of a program, they should be used simultaneously rather than one after the other. This is because cues amplify the effects of contemporaneous incentives, but crowd-out the spillover effects of past incentives.

## 1.6 Concluding Remarks

### 1.6.1 Recap

This paper proposes a model of inattentive choice, focusing on two factors determining attention: cues and rehearsal.

Applied to repeated actions, the model provides an attention-based theory of “good” habits and provides a unifying explanation of many recent empirical findings in domains of behavior ranging from exercise to residential energy use. The model is distinguished from theories of habit-forming *preferences* (Becker and Murphy, 1988) by its predictions about how 1) cues affect current and future behavior 2) how cues can amplify the effects of temporary shocks to routines and 3) how cues can also diminish the effects of temporary shocks to routines. Consistent with these predictions, the first experiment in this paper uses a three-week, real-effort task to test and confirm the distinguishing prediction that repeat performance and reminders are substitutes.

Applied to tasks that must be completed by a deadline, the model identifies when the likelihood of being attentive to the task will decay with time, which then leads naive decision makers to be hurt by longer deadlines. But at the same time, the model also generates new comparative statics about how changes in cues can break the time-decay of attention and therefore eliminate the “perverse” effect of longer deadlines. The second experiment reported in this paper replicates the finding that longer deadlines can lead to

lower completion rates but, consistent with the new comparative statics about cue effects, shows that a small set of reminders can eliminate this effect.

Finally, I apply the model to study market interactions between sophisticated firms and inattentive consumers. Building on the results about tasks with deadlines, I show how firms will use sales tactics such as consumer rebates and automatic renewal billing to facilitate deceptively attractive and socially inefficient transactions with consumers. Building on the results about repeated actions, I use the model to study a firm or organization's optimal policy of *reminder advertising* to inattentive consumers who must take some repeated action. In contrast to models of informative advertising, the optimal messaging intensity does not converge to zero in the long run and may be non-monotonic in consumers' preferences for taking the action.

Taken together, the theoretical and experimental results show that 1) time-varying attention can explain a broad range of economically important behaviors that are not easily explained by other theories; 2) the effects of payoff-irrelevant cues on economic behaviors can be predicted and studied with theoretical and empirical precision; 3) theoretically motivated variations in cues can be used to empirically distinguish between time-varying attention and other potential drivers of behavior; and 4) consumer inattention may be an important determinant of firms' sales and marketing strategies.

Industry players', policymakers' and academics' growing interest in combining incentives with cue-based interventions to encourage behaviors ranging from preventive health measures to energy conservation—as well as the need to understand the medium- and long-run effects of these (sometimes short-run) programs—makes the developments in this paper especially pertinent. The theoretical model and experimental frameworks developed in this paper can provide useful structure for future empirical work on these questions.

### **1.6.2 Extensions**

A stark and not fully realistic characteristic of the model in this paper is the assumption that all else equal, inattention does not vary with stakes. Some models of bounded rationality



(e.g., Mullainathan 2002) make a similar assumption, while others (e.g., Sims 2003) propose that a limited cognitive resource is allocated optimally toward the highest stake decisions.<sup>67</sup> In this paper, higher stakes may make sophisticated agents more motivated to engage in rehearsal or, as in Appendix A.1.2, to invest in reminder technologies. An intuition that is absent from the model, however, is that even in the absence of all cues, a person would be more attentive to higher stakes tasks. At the same time, assuming that attention allocation is completely “rational” in the sense of Sims (2003) is discordant with the fact that many people are naive about how attentive they will be in the future. A completely “rational” approach to attention also seems to be at odds with the notion that sometimes people just pay attention to the “wrong” things. One avenue for future theoretical and experimental work is to investigate how the types of inattentive choices modeled in this paper might depend on the stakes of the decision.

A second way in which the model in this paper could be extended is by incorporating the possibility that some types of contextual cues derive their strength from repeated pairing with a certain action, as in the work of Laibson (2001) and Bernheim and Rangel (2004) on addiction. Interventions to improve medication adherence, for example, recommend associating the action of taking medicine with an event-based cue such as breakfast (Insel et al., 2013). Endogenously formed associations between context and actions may further help explain why some routines, such as brushing one’s teeth every evening, are *so* stable. Relatedly, people may also become desensitized to cues that they receive often but typically don’t pair with the action (Rankin et al., 2009). The sensitization and desensitization considerations may play an important role in the design of optimal communication strategies such as the ones analyzed in Section 1.5.2, and limit the extent to which communications can be used to increase task completion, as in the examples studied in Section 1.4. The experimental design described in Section 1.3.2 could be extended to study desensitization to reminders, as well as the sensitization to event-based cues.

---

<sup>67</sup>Psychologists recognize that attention has both a “goal-driven”/“top-down” component and a “stimulus-driven”/“bottom-up” component (Yantis, 1998)

Finally, future work should broaden the choice set beyond the simple decisions studied in this paper. A key concept to theoretically and experimentally explore in this more general framework is the hypothesis that making some actions more top of mind might crowd out attention to other actions. Consider, for example, a driver who must choose between left, right, or straight at an intersection, and suppose that left is the correct choice on the driver's usual route to work. In the rare and atypical instances in which the driver must, instead, go straight, he may still go with the routine choice of left when he is inattentive. This example illustrates the intuition that rehearsal of one action may make it very mentally accessible, but at the expense of making other actions less so.<sup>68</sup> Extending the model to larger choice sets and, with that, formalizing the nuances of how attention is allocated to the different alternatives is a modeling challenge that may shed light on the formation and power of defaults, and may generate new insights about how firms compete through advertising policies.

---

<sup>68</sup>At the same time, there seem to be some *a priori* default choices such as *not* taking the daily medication or *not* turning off the lights whose consideration is much less likely to be crowded out by the increased accessibility of the alternative(s). My speculation is that the option of *not* performing an action is salient whenever execution of the action requires a certain degree of cognitive monitoring.

## Chapter 2

# The Lightbulb Paradox: Evidence from Randomized Experiments<sup>1</sup>

### 2.1 Introduction

It has long been suggested that consumers may be imperfectly informed about or inattentive to energy costs when they buy energy using durables such as cars, air conditioners, and lightbulbs.<sup>2</sup> This suggestion is supported by recent empirical evidence from other domains: people are inattentive to "add-ons" or ancillary product costs such as sales taxes (Chetty, Looney, and Kroft 2009),<sup>3</sup> shipping and handling charges (Hosseini and Morgan 2006), and out-of-pocket insurance costs (Abaluck and Gruber 2011). Because American households spent \$325 billion on gasoline and another \$245 billion on electricity, natural gas, and heating oil in 2011 (BLS 2013), even small inefficiencies can aggregate to substantial losses.

In theory, the first best policy to address imperfect information and inattention would be an idealized information provision technology that is both costless and "powerful", by

---

<sup>1</sup>Co-authored with Hunt Allcott

<sup>2</sup>See Anderson and Claxton (1982), Blumstein *et al.* (1980), Jaffe and Stavins (1994), Sanstad and Howarth (1994), Gillingham and Palmer (2013), and many others.

<sup>3</sup>Relatedly, Finkelstein (2009) and Cabral and Hoxby (2012) show how salience varies across different tax collection methods.

which we mean that all treated consumers would become fully informed about and attentive to energy costs and other product attributes. In practice, the U.S. and other countries have energy use disclosure requirements such as fuel economy labels on new cars and "yellow tags" on home appliances. There is little evidence on how effectively these programs inform consumers or how they affect purchases.

In addition to information disclosure, policymakers also have a broad set of second best corrective policies such as subsidies and standards for energy efficient autos, appliances, and buildings. Along with externalities and the so-called "landlord-tenant" agency problem, imperfect information and inattention are key potential justifications in both academic papers<sup>4</sup> and government regulatory impact analyses.<sup>5</sup> Evaluating these policies is important, as they are costly: fuel economy standards, appliance energy efficiency standards, "demand-side management" programs run by electric and gas utilities, and weatherization subsidies cost \$17 billion each year (Allcott and Greenstone 2012).

This paper focuses on one the lightbulb market, a particularly compelling case study of consumer behavior and public policy. Do people know how much electricity a traditional incandescent uses relative to an energy efficient compact fluorescent lightbulb (CFL)? Do we pay as much attention to this additional cost as we do to purchase prices, or are we inattentive, like consumers in Gabaix and Laibson (2006)? How should the government intervene, if at all? Regulated and government-run electric utilities spent \$252 million subsidizing and otherwise promoting energy efficient compact fluorescent lightbulbs (CFLs) in the U.S. in 2010 (DOE 2010). Furthermore, the Energy Independence and Security Act of 2007 sets minimum efficiency standards that ban traditional incandescent lightbulbs

---

<sup>4</sup>Among other references, see Gillingham and Palmer (2013), Fischer, Harrington, and Parry (2007), and Parry, Evans, and Oates (2010). The latter paper, for example, focuses on two market failures that could justify energy efficiency standards: externalities and what they call "misperceptions market failures."

<sup>5</sup>See, for example, the Regulatory Impact Analysis for the increase in the Corporate Average Fuel Economy standard for 2012 to 2016. The analysis argues that even without counting the externality reductions, the regulation increases consumer welfare, perhaps because consumers have incorrect "perceptions" of the value of fuel economy (NHTSA 2010, page 2). See also the Regulatory Impact Statement for Australia's ban on energy inefficient lightbulbs (DEWHA 2008, page vii), which argues that "information failures" and consumer cognitive costs help to justify that policy.

between 2012 and 2014 and will be tightened further in 2020.

This ban on traditional incandescents has generated vigorous debate. Many consumers dislike CFLs because they are inferior on several dimensions, and opponents suggest that the regulation is "an example of over-reaching government intrusion into our lives" (Formisano 2008). The government's economic impact analysis (DOE 2009) and other studies (NRDC 2011) argue for the policy by showing that consumers will enjoy billions of dollars in annual cost savings. Of course, such private cost savings could only reflect welfare gains in the presence of imperfect information, inattention, or some other market failure. Argentina, Australia, Brazil, Canada, China, Cuba, the European Union, Israel, Malaysia, Russia, and Switzerland have also banned some or all incandescent light bulbs.

We combine two randomized information provision experiments with a formal model of optimal policy to answer two research questions. First, *how much can information provision affect demand for energy efficient lightbulbs?* Second, *if powerful information provision is costly or infeasible, do subsidies and minimum standards increase welfare as second best solutions to imperfect information and inattention?*

We begin by presenting the results of two randomized control trials (RCTs) that measure the effects of energy cost information on lightbulb purchases. The first is an "artefactual field experiment" (Levitt and List 2009) using Time-Sharing Experiments for the Social Sciences (TESS). This is a high-quality computer-based survey platform which has been used by a number of economists, including Allcott (2013), Fong and Luttmer (2009), Heiss, McFadden, and Winter (2007), Newell and Siikamaki (2013), Rabin and Weizsacker (2009), and others. We gave consumers a \$10 shopping budget and asked them to make a series of choices between CFLs and incandescents in a multiple price list format. We then gave the treatment group information about lightbulb energy costs and lifetimes. After this informational intervention, we again asked all consumers to choose between CFLs and incandescents. This design allows us to infer the joint distribution of demand in the baseline and "informed" states, which is crucial for policy analysis. The experiment was incentive compatible: one of the choices was randomly selected to be the consumer's "official purchase," and the

consumer received those lightbulbs and kept the remainder of his or her shopping budget. The informational intervention increased average willingness to pay for CFLs by \$2.32, and CFL market share at market prices increased by about 12 percentage points.

The second RCT is a "framed field experiment" (Levitt and List 2009) with a large home improvement retailer. Our staff intercepted shoppers, used an iPad to deliver information about energy costs and bulb lifetimes to the treatment group, and then gave coupons with randomly assigned CFL subsidies. While a 20 percent subsidy increased CFL market share by about 10 percentage points, the informational intervention had statistically zero effect. We can bound the information effect as less than the effect of a 12 percent subsidy with 90 percent confidence.

The second half of the paper uses the experimental results to analyze the welfare effects of second best subsidies and standards, given that our informational interventions are not realistically feasible at large scale. In doing this, we follow the approach of Chetty, Looney, and Kroft (2009) in assuming that our treatment groups made informed and otherwise optimal decisions, meaning that our treatment effects measure the magnitude of bias from imperfect information and inattention. Qualitatively, we believe that this assumption is best viewed as an approximation, and we present evidence to evaluate it throughout the paper. In order to ensure that the assumption would be particularly reasonable, we designed the two interventions to fully inform consumers and draw attention to energy costs and other attributes, while minimizing other possible effects. For example, the interventions included no information or cues related to environmental externalities, other social costs of energy use, or social norms. We also took a series of steps to minimize experimenter demand effects. Furthermore, we delivered information through different channels and quizzed consumers on comprehension in the TESS experiment, thus ensuring that consumers understood and attended to the information.

In order to use the empirical results for policy analysis, we formalize a simple theoretical framework that clarifies the "internality rationale" for energy efficiency policy. Consumers make a discrete choice, which in our application is between an incandescent and a CFL.

Some consumers, however, may misjudge the true difference in utility they would experience from the two products; we label the dollar value of this potential mistake the "internality." The internality is directly analogous to an externality: it is a wedge between willingness to pay and social welfare, and it may be heterogeneous across consumers. The policymaker has two instruments: an "internality tax" (in our example, a CFL subsidy) with lump-sum recycling and a ban on the "sin good" (in our example, a ban on traditional incandescents). Just as Diamond (1973) shows that the optimal externality tax equals the average marginal externality, the optimal internality tax equals the average marginal internality. The welfare effect of the ban is the sum of the true utility experienced by the set of consumers who would buy the banned good if they were allowed to do so. Crucially, this average marginal internality is a sufficient statistic in the sense of Chetty (2009): the underlying "structural" model of the bias and any heterogeneity within the set of marginal consumers are both irrelevant for evaluating the welfare effects of a subsidy or ban.

In the context of our theoretical model, the TESS experiment results suggest that the optimal subsidy is approximately \$3 per 60-Watt equivalent CFL. This is slightly larger than typical CFL subsidies offered by many electric utilities. However, we also observe a large group of consumers who purchase incandescents at baseline and are still willing to pay substantially more for incandescents after the informational intervention. Banning incandescents imposes welfare losses on this population that outweigh the gains to apparently-biased consumers who had weaker preferences for the incandescent. This implies that in our model, imperfect information and inattention by themselves do not justify a ban on traditional incandescents.

The simpler design of the in-store experiment identifies only the slope of demand and the effect of information on quantity demanded. However, we show how these two parameters can be used to derive a first-order approximation to the optimal subsidy. Intuitively, the average marginal internality is the price change that would have the same effect on demand as the informational intervention. Given that the intervention had statistically zero effect, we cannot reject that the optimal subsidy is zero. Our formula bounds the optimal subsidy

for a 60-Watt equivalent CFL between negative 30 cents and positive 35 cents per CFL with 90 percent confidence. Given the difference in the population and the experimental setting, it does not surprise us that the effects differ from the TESS experiment. This result only strengthens the qualitative conclusion that the internalities we consider are not large enough to solely justify a ban in our model.

The paper makes three central contributions. First, our two experiments are a "proof of concept" for how large-sample randomized control trials can be used to test the effects of energy use information on durable goods purchases.<sup>6</sup> The dearth of evidence in this context is especially remarkable given the large literature on the effects of information disclosure on consumer choice in other domains, including Choi, Laibson, and Madrian (2010) and Duarte and Hastings (2012) on financial choices, Greenstone, Oyer, and Vissing-Jorgensen (2006) on securities, Bhargava and Manoli (2013) on takeup of social programs, Jin and Sorensen (2006), Kling *et al.* (2012), and Scanlon *et al.* (2002) on health insurance plans, Pope (2009) on hospitals, Bollinger, Leslie, and Sorensen (2011) and Luo *et al.* (2012) on health and nutrition, Dupas (2011) on HIV risk, Figlio and Lucas (2004) and Hastings and Weinstein (2008) on school choice, and many others.

Second, there is a growing empirical literature on whether consumers of durable goods "undervalue" energy costs relative to upfront prices, including Allcott (2013), Allcott and Wozny (2013), Busse, Knittel, and Zettelmeyer (2013), Dubin and McFadden (1984), Goldberg (1998), Hassett and Metcalf (1995), Hausman (1979), Metcalf and Hassett (1999), Sallee, West, and Fan (2009), and many others. Imperfect information and inattention are two of the factors that could cause undervaluation. Most of the previous literature has tested

---

<sup>6</sup>There are some related studies that differ from our experiments on one or more dimensions. Kallbekken, Saelen, and Hermansen (2013) study energy information disclosure at six retail stores in Norway, comparing purchases to a non-randomly selected control group. Anderson and Claxton (1982) study energy information labels with 12 stores assigned to treatment groups and six to control. There are a number of studies that randomly assign information disclosure across individual experimental subjects and study effects on stated preferences in hypothetical choices, including Newell and Siikamaki (2013) and Ward, Clark, Jensen, Yen, and Russell (2011). Deutsch (2010a, 2010b) studies information disclosure to online shoppers, measuring what products they click on and what products they put in online shopping carts, but he does not observe actual purchases. Houde (2012) uses quasi-experimental variation with a structural demand model to estimate how the Energy Star label affects consumer welfare, while Herberich, List, and Price (2011) and Toledo (2013) study how prices and social norm information affect CFL purchases.



for undervaluation by (essentially) comparing price elasticities to energy cost elasticities using variation in purchase prices and energy costs. Our approach is innovative in this literature in that we instead test for undervaluation using experimentally-induced variation in information about and salience of energy costs, as suggested by Chetty, Looney, and Kroft (2009) and DellaVigna (2009). Aside from allowing for more highly-credible identification via randomized control experiments, this approach also isolates the potential effects of imperfect information and inattention from other potential mechanisms such as present bias, which should be unaffected by our informational interventions. Our results are qualitatively consistent with several of the above papers in suggesting that internalities are small in the particular markets that have been studied, and that corrective subsidies and standards may be stronger in these contexts than can be justified by internalities alone.

Third, we provide an example of how techniques from public economics can be combined with psychologically-motivated experiments to provide insight into important public policies. Related analyses include Bernheim and Rangel (2004), Chetty, Looney, and Kroft (2009), Gruber and Koszegi (2004), Gul and Pesendorfer (2007), O'Donoghue and Rabin (2006), Baicker, Mullainathan, and Schwartzstein (2013), and Mullainathan, Schwartzstein, and Congdon (2012) who study taxes when consumers are present biased or otherwise make mistakes. There are also several analyses of energy taxes, energy efficiency standards, or subsidies for energy efficient goods when consumers misoptimize, including Allcott, Mullainathan, and Taubinsky (2013), Heutel (2011), Fischer, Harrington, and Parry, and Parry Evans, and Oates (2010). Our theoretical framework is straightforward, and closely follows Allcott, Mullainathan, and Taubinsky (2013), Baicker, Mullainathan, and Schwartzstein (2013), and Mullainathan, Schwartzstein, and Congdon (2012). This paper is distinguished from most existing work in "behavioral public economics" in that it combines a theoretical framework with parameters from randomized experiments to derive optimal policy.

Section 2 gives more background on lightbulbs and related policies. Sections 3 and 4 present the TESS and in-store experiments, respectively. Section 5 lays out our theoretical framework and derives optimal policies and welfare formulas. Section 6 contains the policy

evaluations, and Section 7 concludes.

## 2.2 Background

### 2.2.1 "The Lightbulb Paradox"

Lightbulbs are a canonical example of the "Energy Paradox" (Jaffe and Stavins 1994): the low adoption of energy efficient technologies despite potentially large savings. Compared to standard incandescents, compact fluorescent lightbulbs (CFLs) typically last eight times longer and use four times less electricity. Although CFLs cost several dollars to purchase, compared to a dollar or less for incandescents, using a CFL saves about \$5 each year once the costs of electricity and replacement bulbs are included. Despite this cost advantage, only 28 percent of residential sockets that could hold CFLs in 2010 actually had them (DOE 2010). In that year, using incandescents instead of CFLs cost US households \$15 billion.<sup>7</sup> Although one lightbulb is inexpensive and by itself uses little electricity, this aggregate figure makes it difficult to argue that the lightbulb market is unimportant, especially when viewed as a case study of issues relevant to the broader class of energy-using durables.

Of course, CFLs and incandescents are differentiated products: many consumers do not like CFLs because the light quality is different, they sometimes flicker, they take time to reach full brightness, and they must be properly disposed of because they contain mercury. While about 60 percent of Americans report in recent surveys that they are "excited" about the lightbulb efficiency standards, about 30 percent say that they are "worried" because they "prefer using traditional lightbulbs" (Sylvania 2012). As Jaffe and Stavins (1994), Allcott and Greenstone (2012), and many others have pointed out, these kinds of non-financial utility costs from energy efficiency are important potential explanations for the apparent

---

<sup>7</sup>Throughout the paper, we assume that incandescents and CFLs last an average of 1000 and 8000 hours, respectively. (To receive the Energy Star rating, a CFL model must last a median of 8000 hours in official tests. Of course, a given consumer may experience varying results.) The national average electricity price is \$0.10 per kilowatt-hour. Our cost estimate of \$15 billion is equal to 5.8 billion residential sockets (DOE 2012), times the 80 percent of sockets that can accommodate CFLs (DOE 2010) minus the actual "socket share" of 28 percent (DOE 2010), times \$5 per socket per year.

"Energy Paradox." Our framework is very clear in allowing these utility differences, and our results indeed show that many consumers strongly prefer incandescents even after the informational interventions.

The U.S. lighting efficiency standards do not require CFLs, nor do they ban incandescents. Instead, they set a maximum energy use per unit of light output. Along with CFLs, light-emitting diodes (LEDs) and high-efficiency halogen bulbs also comply with the standard. We focus on the choice between CFLs and incandescents because these are *by far* the most important current technologies. In 2012, about 1.5 billion incandescents and 300 million CFLs were purchased, compared to only 23 million LEDs (Energy Star 2013). Our quantitative welfare calculations would change in the future if LEDs become a relevant part of the choice set. However, it seems plausible that the qualitative lessons about imperfect information and inattention from CFLs also apply to LEDs, given that LEDs also have high purchase prices, long lifetimes, and large energy cost savings relative to both incandescents and CFLs.

### **2.2.2 Economic Reasons for Standards and Subsidies**

Governments often intervene in markets to subsidize goods or ban bads. Review articles such as Allcott and Greenstone (2012), Gillingham and Palmer (2013), Jaffe and Stavins (1994), and many others discuss the economic reasons in the context of energy-using durables. One potential reason for such policies is externalities. In the case of lightbulbs, one might think that electricity prices are below social cost due to unpriced externalities from climate change, and banning energy inefficient lightbulbs is a welfare-improving second best policy in the absence of a price on carbon dioxide emissions. However, two other distortions that imply that the marginal price of electricity used for residential lightbulbs could actually be *above* social marginal cost. First, retailers typically include much of fixed distribution costs in marginal prices, as Borenstein and Davis (2012) and Davis and Muehlegger (2010) show for natural gas. Second, most residential customers are charged time-invariant prices instead of the optimal peak-load prices, which are lower at night and higher during the day. If lightbulbs are more likely than to be used at night, they thus use electricity which is

underpriced. This suggests that if the primary distortion is mispriced residential electricity, it could actually be optimal to *subsidize* incandescents.<sup>8</sup>

Policymakers might also subsidize new or emerging products to help correct for uninternalized spillovers from research and development or consumer learning. However, 70 percent of consumers report having at least one CFL in their home, compared to 80 percent who report having at least one incandescent (Sylvania 2012), so the technology is available and the vast majority of consumers already have experience with it.

Agency problems in real estate markets could also justify subsidies and standards. For example, home buyers and renters cannot costlessly observe energy efficiency, which reduces the incentive of sellers and landlords to install energy efficient capital stock. Empirical studies by Davis (2010) and Gillingham, Harding, and Rapson (2012) provide some evidence of this, and we are able to provide some evidence from the TESS experiment as well. A final set of inefficiencies is "internalities," which we define as choices that don't maximize the decision maker's own welfare. Present bias over cash flows could be one such externality: if consumers underweight future energy cost savings, they would be less energy efficient than their long-run preferences would dictate, and sophisticates would demand commitment devices to make their future selves buy CFLs and hybrid cars. However, as Andreoni and Sprenger (2012) and others have pointed out, agents in most models are present biased over *consumption*, and most consumers have enough liquidity that paying the incremental few dollars for a CFL does not immediately affect consumption.

Our paper focuses on a particular class of externalities, imperfect information and inattention. In the absence of our results, what would empirical estimates from other contexts suggest could be the magnitudes of these biases? Abaluck and Gruber (2011) find that consumers are five times more responsive to insurance plan premiums than to

---

<sup>8</sup>California is a particularly stark example. Regulations encouraging low-carbon electricity generation mean that the carbon content of electricity consumed there is extremely low relative to other states, so the downward distortion to electricity prices from the lack of a carbon tax is particularly small. Meanwhile, residential electricity tariffs with sharply increasing block prices distort marginal prices upward. Despite the fact that these two forces significantly weaken or reverse the argument that underpriced electricity justifies energy efficiency policies, California has implemented the federal lighting efficiency standards early.

out-of-pocket costs. In their two empirical studies, Chetty, Looney, and Kroft (2009) estimate that consumers are only 35 percent and 6 percent as attentive to sales taxes as they are to product prices. A CFL saves an undiscounted \$36 over its expected life relative to an incandescent. If consumers were (hypothetically) as inattentive to these savings as they are to sales taxes, these estimates suggest that they could undervalue the CFL's energy savings by \$23 to \$34. This dwarfs the typical difference in purchase prices between CFLs and incandescents, and it suggests that our informational interventions could have massive impacts on demand.

In summary, while there are other market failures that could justify subsidies and standards for lightbulbs, we designed this study to focus on imperfect information and inattention because results from other literatures suggested that these two distortions could be large, while other market failures appear to be less relevant.

## **2.3 TESS Experiment**

### **2.3.1 Survey Platform**

We implemented the artefactual field experiment through Time-Sharing Experiments for the Social Sciences (TESS). TESS, which is funded by the National Science Foundation, facilitates academic access to KnowledgePanel, an online experimental platform managed by a company called GfK. The platform has been used by a number of economists, including Allcott (2013), Fong and Luttmer (2009), Heiss, McFadden, and Winter (2007), Newell and Siikamaki (2013), and Rabin and Weizsacker (2009).

One reason why economists use TESS is the recruitment process, which generates a sample as close as practically possible to being nationally representative on unobservable characteristics and reduces concern about generalizability. Potential KnowledgePanel participants are randomly selected from the U.S. Postal Service Delivery Sequence File and recruited through a series of mailings in English and Spanish, plus telephone-based follow-up when the address can be matched to a phone number. About 10 percent of

people who are invited actually consent and complete the demographic profile to become KnowledgePanel participants, and there are now approximately 50,000 active panel members. Unrecruited volunteers are not allowed to opt in. Households without computers are given computers in order to complete the studies.

KnowledgePanel participants take an average of two studies per month, and no more than one per week. Of the KnowledgePanel participants who started our study, some were not qualified to continue because their computer audio did not work, which would have prevented them from hearing the audio part of our treatment and control interventions. Of the qualified participants who began the survey, about 3/4 completed it, giving us a final qualified sample size of 1533. Although we were able to negotiate with GfK to require answers on some of the most important parts of the study, GfK's policy is to usually allow participants to refuse. A handful of participants refuse to answer any given question, so the number of observations will vary slightly across regressions.

GfK provides sampling weights which can allow us to match the US population aged 18 and older on gender, age, ethnicity, education, census region, urban or rural location, and whether the household had internet access before recruitment. All statistics presented in the paper are weighted for national representativeness, although weighting observations equally does not substantively change the results.

### **2.3.2 Experimental Design**

The study had four parts: baseline lightbulb choices, the informational intervention, endline lightbulb choices, and a post-experiment survey. About 21 percent of treatment group consumers (226 consumers) were randomly assigned to skip the baseline lightbulb choices and begin directly with the informational intervention; these "Endline-Only" consumers are included in all statistics when possible and are the focus of several robustness checks. Appendix B.1 contains screen shots from each part of the experiment.

Consumers were first shown an introductory screen with the following text:

*In appreciation for your participation in this study, we are giving you a \$10 shopping budget.*

*With this money, we will offer you the chance to buy light bulbs. You must make a purchase with this money. Whatever money you have left over after your purchase, you get to keep. This money will be provided to you as cash-equivalent bonus points that will be awarded to your account.*

*In approximately four to six weeks, GfK will send you the light bulbs you have purchased. Light bulbs are frequently shipped in the mail. There is not much risk of breakage, but if anything does happen, GfK will just ship you a replacement. Even if you don't need light bulbs right now, remember that you can store them and use them in the future.*

*During the study, we will ask you to make 30 decisions between pairs of light bulbs. There will be a first set of 15 decisions, then a break, and then a second set of 15 decisions. After you finish with all 30 decisions in the questionnaire, one of them will be randomly selected as your "official purchase." GfK will ship you the light bulbs that you chose in that official purchase. Since each of your decisions has a chance of being your official purchase, you should think about each decision carefully.*

### **Baseline Lightbulb Choices**

After the introductory screen, consumers were then shown two lightbulb packages. One package contained one Philips 60-Watt equivalent Compact Fluorescent Lightbulb. The other contained four Philips 60-Watt incandescent lightbulbs. The two choices were chosen to be as comparable as possible, except for the CFL vs. incandescent technology. Half of respondents were randomly assigned to see the incandescent on the left, labeled as "Choice A," while the other half saw the incandescent on the right, labeled as "Choice B."

Consumers had the option to "click for detailed product information," and about 19 percent did so. This opened a simple "Detailed Product Information" screen, which included the light output in Lumens, a quantitative measure of light color, energy use in Watts, and other information. Both packages typically sell online for about \$4, although we did not tell consumers these typical prices.

Lower down on this same screen, consumers were asked to make their baseline lightbulb choices: 15 decisions between the same two packages at different relative prices. Decision Number 1 offered Choice A for free and Choice B for \$10. The relative price of Choice A

increased monotonically until Decision Number 15, which offered Choice A for \$10 and Choice B for free. Consumers spent a median of three minutes and zero seconds to complete these first 15 decisions.

We identify consumers' baseline relative willingness to pay (WTP) for the CFL, denoted  $v_i^0$ , using the relative prices at which they switch from preferring CFLs to incandescents. For example, consumers who choose CFLs when both packages cost \$4 but choose incandescents when incandescents are one dollar cheaper are assumed to have  $v_i^0 = \$0.50$ . Eight percent of consumers did not choose monotonically: they chose Choice A at a higher relative price than another decision at which they chose Choice B. These consumers were prompted with the following message: *The Decision Numbers below are organized such that Choice A costs more and more relative to Choice B as you read from top to bottom. Thus, most people will be more likely to purchase Choice A for decisions at the top of the list, and Choice B for decisions at the bottom of the list. Feel free to review your choices and make any changes. Then click NEXT.* After this prompt, 5.3 percent of consumers still chose non-monotonically, and we code their WTP as missing.

Some consumers had "censored" WTPs: they preferred either Choice A or Choice B at all relative prices. These consumers were asked to self-report their WTP. For example, a participant who always preferred Choice A was asked: *Your decisions suggest that you prefer Choice A even when Choice A costs a total of \$10 and Choice B is free. If Choice B continued to be free, how much would Choice A need to cost in order for you to switch to Choice B? (This is purely hypothetical - your answer will not affect any of the prices you are offered.)*

At baseline, five percent of consumers preferred the incandescents by more than \$10, while 19 percent preferred the CFLs by more than \$10. Across all censored consumers, the median absolute value of self-reported relative WTP was \$15. The distribution of self-reports is skewed, with about eight percent of consumers preferring one or the other choice by more than \$40. Because these are self-reports, we wish to be cautious about using them in the analysis, so we instead assume a mean relative WTP of \$15 and -\$15 for top-coded and bottom-coded consumers, respectively. We will demonstrate the sensitivity of the results to this assumed mean censored value. While an assumed mean censored value may seem



unsatisfying, remember that in the absence of this type of experiment, a demand model used to predict the removal of a product from the choice set would typically assume a logit or otherwise parametric functional form for demand.

In theory, if lightbulbs were perishable and consumers did not immediately need one, they would buy the cheapest package instead of revealing the WTP they would have if they did need one. In practice, lightbulbs are easily stored, and we reminded consumers of this fact in the introductory text. In theory, if it were costless to resell the experimental purchase and replace it with a different purchase outside the experiment, consumers who know that the typical retail prices are approximately equal would always buy the cheaper package. In order to avoid making this salient, the experiment website did not include information about the bulbs' typical retail prices. In practice, it seems unlikely that consumers resold the packages that they received. If non-storability and price arbitrage affected some consumers' choices, this would make the demand curve more elastic and the treatment effects less positive. Empirically, however, we see large shares of consumers with relative WTPs that differ substantially from typical market relative prices.

### **Informational Intervention**

Consumers were randomized into three groups: Balanced treatment, Positive treatment, and control. Endline-Only consumers were randomized between the two treatment groups with equal probability. All other consumers were randomized between all three groups with equal probability.

The information treatments were designed to give clear product information while minimizing the possibility that the information would be perceived as biased. The treatments were also designed to closely parallel each other, to minimize the chance that idiosyncratic factors other than the information content could affect purchases. Each treatment had the following structure:

1. Belief elicitation. This elicited prior beliefs over the information to be presented in the two Information Screens.

2. Introductory Screen

3. First Information Screen. This had text plus an illustratory graph, and the text was also read verbatim via an audio recording. The audio recordings are available as part of the Online Supplementary Materials. At the bottom of the information screen, there was a "quiz" on a key fact.

4. Second Information Screen. This paralleled the First Information Screen.

The central difference between the three conditions was the content of their two Information Screens. The order of these two screens was randomly assigned with equal probability.

We took two steps to make sure that all consumers understood the treatment. First, we used multiple channels to convey information: text, graphical, and audio. This means that people who process information in different ways had a higher chance of internalizing the information. Second, the quiz forced respondents to internalize the information if they had not done so initially.

**Balanced Treatment** The Balanced treatment Introductory Screen had the following text:

*For this next part of the study, you will have the opportunity to learn more about light bulbs. We will focus on the following two issues:*

1. *Total Costs*

2. *Disposal and Warm-Up Time*

*The discussion of each issue will be followed by a one-question quiz. Please pay close attention to the discussion so that you can correctly answer the quiz question.*

Consumers then advanced to the Total Cost Information Screen and the Disposal and the Warm-Up Information Screen, in randomized order. The Total Cost Information Screen explained that CFLs both last longer and use less electricity and translates these differences into dollar amounts. The bottom line was:

*Thus, for eight years of light, the total costs to purchase bulbs and electricity would be:*

- *\$56 for incandescents: \$8 for the bulbs plus \$48 for electricity.*
- *\$16 for a CFL: \$4 for the bulbs plus \$12 for electricity.*

The quiz question at the bottom of the screen was: *For eight years of light, how much larger are the total costs (for bulbs plus electricity) for 60-Watt incandescents as compared to their CFL equivalents?* The correct answer could be inferred from the information on the screen: \$56 for incandescents - \$16 for CFLs = \$40. Sixty-four percent of consumers correctly put \$40. Those who did not were prompted: *That is not the correct answer. Please try again.* After this prompt, 73 percent of consumers had typed \$40. The remaining consumers were prompted: *The total costs for eight years of light are \$16 for CFLs and \$56 for incandescents. Therefore, the incandescents cost \$40 more. You may type the number 40 into the answer box.* By this point, 89 percent of consumers had correctly typed \$40. This documents that the vast majority of consumers understood at least some part of the information. Consumers spent a median of two minutes and 12 seconds to read the Total Cost Information Screen and complete the quiz question.

The Disposal and Warm-Up Information Screen was designed to present information about ways in which CFLs may not be preferred to incandescents. It paralleled exactly the Total Cost Information Screen, beginning with belief elicitation, and then continuing to an Information Screen with text of similar length, a graph, and a quiz question at the bottom. The Disposal and Warm-Up Information Screen explained that *"because CFLs contain mercury, it is recommended that they be properly recycled instead of disposed of in regular household trash."* It also explained that *"after the light switch is turned on, CFLs take longer to warm up than incandescents."* We included this information to reduce the probability of experimenter demand effects, through which consumers might think that the experimenter wanted them to purchase the CFL, potentially causing their endline choices to differ from true preferences.

**Control** The control intervention was designed to exactly parallel the treatment interventions, but with information that should not affect relative WTP for CFLs vs. incandescents. One screen presented the number of lightbulbs installed in residential, commercial, and industrial buildings in the United States. The other screen detailed trends in total lightbulb sales between 2000 and 2009.

**Positive Treatment** The Positive treatment was designed to inform consumers about the benefits of the CFL in terms of lifetime and lower energy costs, without presenting information about ways that the incandescent might be preferred to the CFL. This more closely parallels the intervention in the in-store RCT described in the next section. To implement this while keeping the intervention the same length, we combined the Total Cost Information Screen with a random draw of one of the two control screens.

### **Endline Lightbulb Choices**

The endline lightbulb choice screen was analogous to the baseline screen. Consumers spent a median of one minute and 20 seconds to complete these final 15 decisions. We determine endline relative WTP  $v_i^1$  in the same way as above.

### **2.3.3 Data**

Column 1 of Table 2.1 presents descriptive statistics. Liberal is self-reported political ideology, originally on a seven-point scale, normalized to mean zero and standard deviation one, with larger numbers indicating more liberal. Party is self-reported political affiliation, similarly normalized from an original seven-point scale, with larger numbers indicating more strongly Democratic. Environmentalist is the consumer's answer to the question, "Would you describe yourself as an environmentalist?" Conserve Energy is an indicator for whether the consumer reports having taken steps to conserve energy in the past twelve months. Homeowner is a binary indicator variable for whether the consumer owns his or

her home instead of rents. These questions were asked when the participant first entered KnowledgePanel, not as part of our experiment.

**Table 2.1:** *Descriptive Statistics and Balance for TESS Experiment*

	(1)	(2)	(3)
<b>Individual Characteristics</b>	Population Mean	Treatment - Control Difference	Positive - Balanced Treatment Difference
Baseline Relative Willingness-to-Pay for CFL (\$)	2.9 (7.1)	0.20 (0.52)	-0.25 (0.70)
Household Income (\$000s)	70.9 (51.8)	-2.86 (3.92)	-3.79 (3.86)
Education (Years)	13.8 (2.5)	-0.04 (0.18)	0.18 (0.21)
Age	46.7 (16.9)	0.26 (1.26)	0.22 (1.45)
Male	0.48 (0.50)	-0.007 (0.035)	-0.009 (0.040)
Liberal	0.00 (1.00)	0.056 (0.076)	-0.005 (0.084)
Party	0.00 (1.00)	0.080 (0.072)	0.078 (0.082)
Environmentalism	0.30 (0.32)	-0.024 (0.023)	0.019 (0.026)
Conserve Energy	0.55 (0.50)	0.008 (0.036)	0.032 (0.041)
Homeowner	0.70 (0.46)	0.022 (0.035)	-0.012 (0.038)
F-Test p-Value		0.848	0.995

*Notes:* Column 1 presents means of individual characteristics in the TESS experiment population, with standard deviations in parenthesis. Column 2 presents differences in means between the treatment groups and control, while column 3 presents differences in means between Positive and Balanced treatment groups. Columns 2 and 3 have robust standard errors in parenthesis. \*, \*\*, \*\*\*: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

Column 2 presents the difference in means between consumers in either of the two treatment groups vs. control. Column 3 presents the difference in means between the Positive and Balanced treatment groups. All 20 t-tests fail to reject equality, as do the joint F-tests of all characteristics. Like all reported results, these are weighted for national representativeness, although the unweighted groups are also balanced on all characteristics.

Appendix Table B2.1 presents correlations between baseline WTP and observable characteristics. Men, democrats, environmentalists, those who report having taken steps to conserve energy, and those with higher discount factors have higher demand for CFLs. (The discount factors are the  $\delta$  parameter in a  $\beta, \delta$  model of present bias, as calibrated from hypothetical intertemporal tradeoffs in the post-experiment survey.) These correlations conform to our intuition and build further confidence that the differences in WTP are meaningful. However, renters and more present-biased (lower  $\beta$ ) consumers do not have lower WTP for CFLs conditional on other observables. This provides no support for the hypotheses that agency problems and present bias play a role in lightbulb decisions.

### 2.3.4 Empirical Strategy and Results

Figure 2.1 shows the baseline and endline demand curve for CFLs. The control endline demand curve sits almost directly on top of the baseline demand curve, implying that the control interventions had little effect on relative demand for CFLs vs. incandescents. The treatment endline curve is shifted outwards, reflecting an increase in demand for the CFL. At equal prices, which approximately reflects market conditions, 76 percent of the treatment group chooses the CFL, against 65 percent of control.

Figure 2.2 presents a histogram of the within-subject changes in WTP between baseline and endline. About 90 percent of control group consumers either have exactly the same WTP or change by \$2 or less. In treatment, there is a mass to the right of the figure, with 36 percent of people increasing WTP by between \$1 and \$10.

Denote  $X_i$  as participant  $i$ 's vector of characteristics from Table 1, and  $T_i$  as an indicator for whether the household is in either of the two treatment groups. We estimate the average treatment effects of the informational interventions on endline willingness-to-pay  $v_i^1$  using OLS with robust standard errors:

$$v_i^1 = \tau T_i + \gamma X_i + \varepsilon_i \tag{2.1}$$

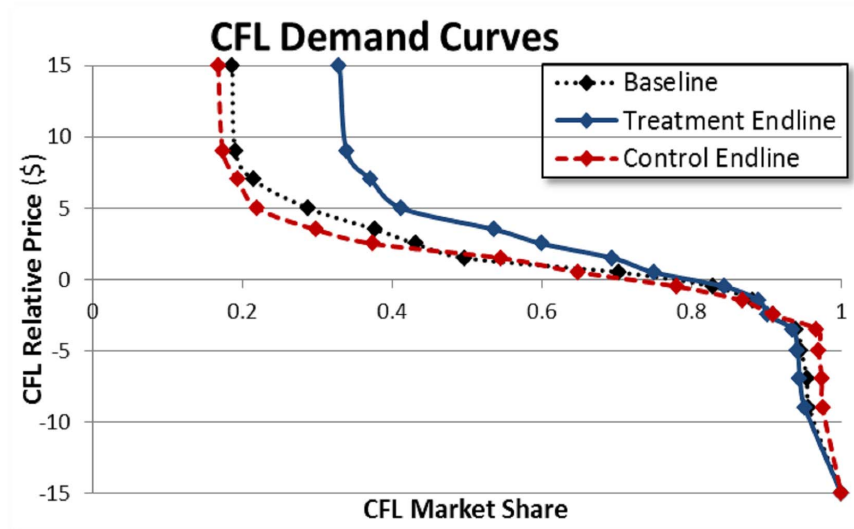


Figure 2.1: CFL Demand Curves

Table 2.2 presents the results. Column 1 presents the unconditional difference in means, excluding the Endline-Only group for comparability with other columns that include baseline WTP. Column 2 adds the control for baseline WTP  $v_i^0$ . Column 3 is the exact specification from Equation (2.1), including individual characteristics. The sample size decreases in column 3 because at least one  $X$  characteristic is missing for 15 consumers, but the effects do not change statistically. In column 3, the informational intervention caused consumers' WTP for the CFL to increase by an average of about \$2.32.

One potential concern is that treatment group consumers might wish to be internally consistent in their choices between baseline and endline (Falk and Zimmermann 2012). This could cause endline choices to be biased towards the baseline, unlike an experimental design that did not require consumers to state baseline choices. This in turn would bias treatment effects toward zero. The Endline-Only treatment group was included to test this. Column 4 includes only the Endline-Only and control groups, excluding the treatment group that made baseline choices. The estimates should be compared against Column 1, which similarly does not control for baseline WTP. The point estimate is not statistically

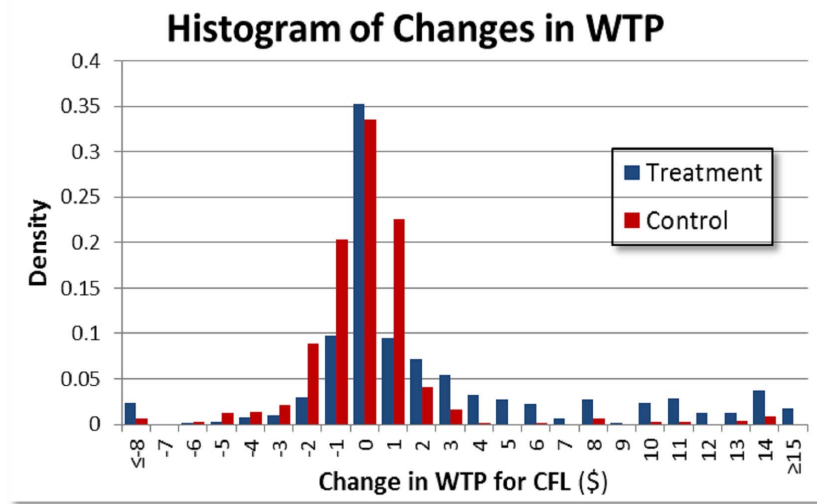


Figure 2.2: Histogram of Changes in WTP

different, although it is lower by \$0.46 per package.<sup>9</sup>

Top-coding and bottom-coding of WTP mechanically influence the treatment effect. Consumers with baseline WTP equal to the maximum cannot reveal a post-treatment increase in WTP, and any consumers with baseline WTP equal to the minimum could not reveal a decrease in WTP. Because the treatment tends to increase WTP, the former effect should dominate, and the average treatment effect should be understated. Column 5 excludes consumers with top-coded or bottom-coded baseline WTP of  $v_i^0 = \$15$  or  $v_i^0 = -\$15$ . The estimated effect increases to \$3.23.<sup>10</sup>

Relatedly, the assumed mean censored value of \$15 caps the increase in WTP that any consumer can reveal. Since a larger share of endline WTP is top-coded in treatment relative

<sup>9</sup>Two other tests confirm why this result holds. First, average post-treatment WTP does not differ statistically between the Endline-Only group and rest of the treatment group. Second, the shape of demand does not differ between these two groups: Appendix Figure B2.1 shows that the Endline-Only demand curve sits very close to the demand curve for the rest of the treatment group, while control group demand is very different. Statistical tests show that the share of consumers with  $v_i^1 > v^+$  does not differ statistically at any level of  $v^+$ .

<sup>10</sup>This increase is consistent with Figure 2.4, which we discuss later. The figure shows that consumers with top coded baseline relative WTP (plus the small group with WTP of \$9) have close to zero conditional average treatment effect, precisely because there is no way for them to increase their WTP in the multiple price list. This illustrates graphically why excluding these consumers increases the treatment effect.



**Table 2.2: Effects of TESS Informational Interventions**

	(1)	(2)	(3)	(4)	(5)	(6)
1(Treatment)	2.535 (0.549)***	2.301 (0.358)***	2.324 (0.364)***	2.078 (0.777)***	3.231 (0.364)***	2.138 (0.498)***
Baseline Willingness-to-Pay		0.777 (0.037)***	0.775 (0.037)***		0.934 (0.065)***	0.776 (0.037)***
1(Positive Treatment)						0.396 (0.573)
R2	0.03	0.56	0.57	0.02	0.33	0.57
N	1,203	1,203	1,188	656	919	1,188
Individual Characteristics	No	No	Yes	No	Yes	Yes
Include Endline-Only Group	No	No	No	Yes	No	No
Exclude Max./Min. Baseline WTP	No	No	No	No	Yes	No

*Notes:* This table presents estimates of Equation (2.1). The outcome variable is endline willingness-to-pay for the CFL. Robust standard errors in parenthesis. \*, \*\*, \*\*\*: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

to control (29 percent vs. 16 percent), increasing this assumed value should increase the treatment effect. In unreported regressions where we alternatively assume mean censored values of \$12 (\$20) instead of \$15, the ATE decreases to \$1.99 (increases to \$2.88).

## Demand Effects

With any experiment other than a natural field experiment, one might be worried about demand effects: that participants change their actions to comply with, or perhaps defy, the perceived intent of the study. We could have designed a lengthier or otherwise more complex experiment to obfuscate our objective of estimating the effects of information disclosure, but this would have added to the cost. If demand effects are present, the likely direction would be to increase treatment group post-intervention WTP, i.e. make the treatment effect more positive. Aside from pointing out that the likely sign of the bias would only reinforce our qualitative conclusions, we also address demand effects in three ways.

First, we designed the experiment to include the Balanced treatment group, which

disclosed both positive and negative information about CFLs. Consumers in this group should be less likely to believe that the experimenters were purely trying to persuade them to purchase the CFL. If demand effects play a large role, effects of the Positive treatment should be inflated relative to the Balanced treatment. Column 6 of Table 2.2 includes an indicator for the Positive treatment group, showing that the effects do not differ statistically, and the point estimates are fairly similar. Because the effects do not differ between the two groups, we combine all treated consumers in all other parts of the analysis.

Second, demand effects are less likely if participants cannot identify the intent of the study. The post-experiment survey asked consumers what they thought the intent of the study was. Multiple responses were allowed. Table 3 presents the share of each group that gave each response. The two treatment groups responded similarly, although the Balanced group was more likely to report that the intent of the study was to "understand what features of lightbulbs are most important to people" and less likely to report that the intent was to "test how well people are able to quantify energy costs." Relative to control, both treatment groups were more likely to respond that the intent of the study was to "understand why people buy incandescents vs. CFLs," "test how well people are able to quantify energy costs," "test whether ability to quantify energy costs affects purchases," and "test whether consumer education affects purchases." The control group was more likely to respond that the intent was to "understand the effects of price changes," "measure whether people make consistent purchases in similar situations," and "test whether the number of bulbs in a package affects purchasing patterns." The dispersion of beliefs within groups suggests that there is not one obvious way in which demand effects might act.

Third, if demand effects are present, they should differentially affect people who are more able to detect the intent of the study and are more willing to change their choices given the experimenter's intent. One existing measure of these issues is the Self-Monitoring Scale, a battery of personality questions developed by Snyder (1974). Snyder writes that the scale is designed to identify individuals who "tend to express what they think and feel, rather than mold and tailor their self-presentations and social behavior to fit the situation."

**Table 2.3: Perceived Intent of TESS Study**

	(1)	(2)	(3)
	Control	Balanced Treatment	Positive Treatment
<i>Do you think that the intent of the study was to . . .</i>			
Understand the effect of price changes on purchasing patterns	0.44	0.34	0.37
Measure whether people make consistent purchases in similar situations	0.31	0.25	0.26
Understand why people buy incandescents vs. CFLs	0.31	0.48	0.47
Test how well people are able to quantify energy costs	0.27	0.38	0.46
Test whether ability to quantify energy costs affects purchases of incandescents vs. CFLs	0.33	0.50	0.54
Test whether the number of bulbs in a package affects purchasing patterns	0.37	0.22	0.26
Test whether consumer education affects purchases of incandescents vs. CFLs	0.41	0.60	0.64
Understand what features of lightbulbs are most important to people	0.30	0.41	0.34
Predict the future popularity of incandescents vs. CFLs	0.30	0.34	0.34
None of the above	0.05	0.08	0.05
Number of Respondents	461	545	519

*Notes:* This table presents the share of consumers in each group who responded that the intent of the study was as listed in the leftmost column. Observations are weighted for national representativeness.

From the set of standard Self-Monitoring Scale statements, we took the most relevant six:

- It's important to me to fit in with the group I'm with.
- My behavior often depends on how I feel others wish me to behave.
- My powers of intuition are quite good when it comes to understanding others' emotions and motives.
- My behavior is usually an expression of my true inner feelings, attitudes, and beliefs.
- Once I know what the situation calls for, it's easy for me to regulate my actions accordingly.

- I would NOT change my opinions (or the way I do things) in order to please someone else or win their favor.

At the very end of the post-experiment survey, we asked consumers to respond to each of these six statements on a five-point Likert scale from "Agree" to "Disagree." We normalize responses to each question to mean zero, standard deviation one, and interact each with the treatment indicator while also controlling for lower-order interactions. While the six Self-Monitoring Scale variables are correlated with each other, none is correlated with endline CFL demand or with the treatment effect, nor is a composite of the six. The estimation results can be found in Appendix Table B2.2.

### **Mechanisms**

How much of the treatment effect is coming from changing information sets vs. directing attention to energy costs? The post-experiment survey elicits beliefs over how much less it costs to buy electricity for a CFL vs. incandescents over the typical 8000-hour life of a CFL, at national average electricity prices. The question is similar, but not identical, to the "quiz" question asked of the treatment group, and the correct answer is \$36 (\$48 for the incandescent minus \$12 for the CFL). Column 1 of Table 4 shows that the treatment increases median beliefs: they are \$25 in control and \$13 higher in treatment. Column 2 of Table 2.4 shows that the treatment also substantially reduces the median absolute error, i.e. the absolute value of the difference between the reported belief and \$36. As in the other tables, the exact sample sizes are slightly smaller than the total number of qualified participants because a handful of participants refused to answer. We use median regressions because reported beliefs have high variance, and median regressions are more robust to extreme outliers than OLS. These results suggest that at least part of the treatment effect is from changing information sets.

The post-experiment survey also asks consumers to rate on a scale of 1-10 the importance of energy use, bulb lifetime, warm-up time, and mercury and disposal in their purchase

**Table 2.4: Effects on Beliefs**

	(1)	(2)
	Savings Belief	Belief Error
1(Treatment)	13.0 (3.8)***	-14.0 (1.8)***
Constant	25.0 (3.8)***	34.0 (1.4)***
R2	.	.
N	1,506	1,506

*Notes:* In the post-experiment survey, consumers were asked their beliefs about the dollar value of electricity cost savings from owning CFLs instead of incandescents. Columns 1 and 2 present median regressions of the effects of the informational interventions on these beliefs and the absolute value of the error in these beliefs, respectively. \*, \*\*, \*\*\*: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

decisions. Table 2.5 presents how the treatments affected these ratings. Both Positive and Balanced treatments decreased the stated importance of purchase prices, consistent with consumers re-orienting away from purchase price as a measure of cost. Point estimates suggest that both the Positive and Balanced treatments increased the importance of energy use and that the Positive treatment also increased the importance of bulb lifetimes. These are the only estimates in the entire analysis whose significance level is affected by the weighting: they are not significant in Table 2.5, but (unreported) regressions show that they are statistically significant when weighting all observations equally. The Positive treatment and control groups do not differ on the importance of warm-up time or mercury and disposal, which is to be expected because neither group received information on these two issues. Interestingly, the Balanced treatment decreased the importance of warm-up time. One potential explanation is that consumers had previously believed that CFL warm-up times were longer, and the treatment reduced the importance of this difference between CFLs and incandescents.

**Table 2.5: Effects on Important Factors in Purchase Decision**

	(1)	(2)	(3)	(4)	(5)
	Price	Energy Use	Bulb Lifetime	Warm-Up Time	Mercury and Disposal
1(Balanced Treatment)	-0.864 (0.208)***	0.147 (0.214)	0.023 (0.201)	-0.943 (0.243)***	-0.294 (0.252)
1(Positive Treatment)	-0.552 (0.218)**	0.202 (0.210)	0.249 (0.187)	0.036 (0.231)	-0.089 (0.244)
Constant	7.747 (0.134)***	7.435 (0.145)***	7.760 (0.133)***	5.406 (0.167)***	6.030 (0.178)***
R2	0.02	0.00	0.00	0.02	0.00
N	1,533	1,478	1,512	1,506	1,518

*Notes:* This table reports treatment effects on self-reported importance of different factors in purchase decisions. Robust standard errors in parenthesis. \*, \*\*, \*\*\*: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

## 2.4 In-Store Experiment

### 2.4.1 Experimental Design

Would the effects of information provision be different in a more typical retail setting compared to the TESS platform? To answer this, we partnered with a large home improvement retailer to implement an in-store experiment. Between July and November 2011, three research assistants (RAs) worked in four large "big box" stores, one in Boston, two in New York, and one in Washington, D.C. The RAs approached customers in the stores' "general purpose lighting" areas, which stock incandescents and CFLs that are substitutable for the same uses.<sup>11</sup> They told customers that they were from Harvard University and asked, "Are you interested in answering some quick research questions in exchange for a discount on any lighting you buy today?" Customers who consented were given a brief survey via iPad in which they were asked, among other questions, the most important factors in their lightbulb purchase decision, the number of bulbs they were buying, and the amount of

<sup>11</sup>This includes standard bulbs used for lamps and overhead room lights. Specialty bulbs like Christmas lights and other decorative bulbs, outdoor floodlights, and lights for vanity mirrors are sold in an adjacent aisle.

time each day they expected these lightbulbs to be turned on each day. The survey did not mention electricity costs or discuss any differences between incandescents and CFLs.

The iPad randomized customers into treatment and control groups with equal probability. For the treatment group, the iPad would display the annual energy costs for the bulbs the customer was buying, given his or her estimated daily usage. It also displayed the total energy cost difference over the bulb lifetime and the total user cost, which included energy costs plus purchase prices. Appendix B.3 presents the information treatment screen. The RAs would interpret and discuss the information with the customer, but they were instructed not to advocate for a particular type of bulb and to avoid discussing any other issues unrelated to energy costs, such as mercury content or environmental externalities. The control group did not receive this informational intervention, and the RAs did not discuss energy costs or compare CFLs and incandescents with these customers.

At the end of the survey and potential informational intervention, the RAs gave customers a coupon in appreciation for their time. The iPad randomized respondents into either the Standard Coupon group, which received a coupon for 10 percent off all lightbulbs purchased, or the Rebate Coupon group, which received the same 10 percent coupon plus a second coupon valid for 30 percent off all CFLs purchased. Thus, the Rebate Coupon group had an additional 20 percent discount on all CFLs. For a consumer buying a typical package of 60 Watt bulbs at a cost of \$3.16 per bulb, this maps to an average rebate of \$0.63 per bulb. The coupons had bar codes which were recorded in the retailer's transaction data as the customers submitted them at the register, allowing us to match the iPad data to purchases.

After giving customers their coupons, the RAs would leave the immediate area in order to avoid any potential external pressure on customers' decisions. The RAs would then record additional demographic information on the customer, including approximate age, gender, and ethnicity. The RAs also recorded this information for people who refused. Finally, the RA recorded the total duration of the interaction. The difference between treatment and control had a mean of 3.17 minutes and a median of 3.0 minutes. This measures the amount of time spent discussing the energy cost information and the differences between

incandescents and CFLs.

## **2.4.2 Data**

Of the 1561 people who were approached, 459 refused, while 1102 began the iPad survey. Of these, 13 broke off after the first question, two broke off later, and 1087 were assigned to a treatment group and given a coupon. Column 1 of Table 2.6 presents descriptive statistics for the sample of customers who completed the survey and were given a coupon. Column 2 presents differences between the 474 people who refused or did not complete the survey and the 1087 who completed, using the demographic characteristics recorded for those who refused. People whom the RAs thought were older, male, Asian, and Hispanic were more likely to refuse. Columns 3 and 4 present differences between the information treatment and control groups and between the rebate and standard coupon groups. In one of the 18 t-tests, a characteristic is statistically different with 95 percent confidence: we have slightly fewer people coded as Asian in the information treatment group. F-tests fail to reject that the groups are balanced.



**Table 2.6: Descriptive Statistics and Balance for In-Store Experiment**

	(1)	(2)	(3)	(4)
	Experimental Sample Mean	Refused - Sample Difference	Treatment - Control Difference	Rebate - Standard Difference
<b>Individual Characteristics</b>				
Energy an Important Factor in Purchase Decision	0.25 (0.43)		0.009 (0.026)	-0.024 (0.026)
Expected Usage (Minutes/Day)	333 (280)		12.8 (17.0)	2.7 (17.0)
Age	43.8 (11.4)	2.3 (0.6)***	0.7 (0.7)	-0.3 (0.7)
Male	0.66 (0.47)	0.06 (0.03)**	0.009 (0.029)	0.003 (0.029)
African American	0.16 (0.37)	-0.04 (0.02)**	-0.001 (0.022)	-0.008 (0.022)
Asian	0.06 (0.24)	0.04 (0.02)**	-0.030 (0.014)**	0.005 (0.015)
Caucasian	0.66 (0.47)	-0.07 (0.03)**	0.037 (0.029)	-0.005 (0.029)
Hispanic	0.07 (0.25)	0.06 (0.02)***	0.001 (0.015)	0.011 (0.015)
Middle Eastern	0.01 (0.12)	0.01 (0.01)	0.002 (0.013)	0.007 (0.007)
F-Test p-Value		0.00	0.742	0.896
<b>Purchase Decisions</b>				
Purchased Any Lightbulb	0.77 (0.42)		0.011 (0.025)	0.027 (0.025)
Purchased Substitutable Lightbulb	0.73 (0.44)		-0.008 (0.027)	0.011 (0.027)

*Notes:* Column 1 presents means of individual characteristics in the in-store experiment sample, with standard deviations in parenthesis. Column 2 presents differences in recorded demographic characteristics between those who refused or did not complete the survey and the experimental sample. Column 3 presents differences in means between treatment and control groups, while column 4 presents differences in means between the rebate and standard coupon groups. Columns 2, 3, and 4 have robust standard errors in parenthesis. \*, \*\*, \*\*\*: Statistically significant with 90, 95, and 99 percent confidence, respectively.

We restrict our regression sample to the set of consumers that purchase a "substitutable lightbulb," by which we mean either a CFL or any incandescent or halogen that can be replaced with a CFL. The bottom panel of Table 2.6 shows that 77 percent of interview respondents purchased any lightbulb with a coupon, and 73 percent of survey respondents purchased a substitutable lightbulb. While information or rebates theoretically could affect whether or not customers purchase a substitutable lightbulb, t-tests show that in practice the percentages are not significantly different between the groups.

### 2.4.3 Empirical Strategy and Results

We denote  $T_i$  and  $S_i$  as indicator variables for whether customer  $i$  is in the information treatment and rebate groups, respectively.  $X_i$  is the vector of individual-level covariates. We estimate a linear probability model<sup>12</sup> with robust standard errors using the following equation:

$$1(\text{Purchase CFL})_i = \eta S_i + \tau T_i + \alpha X_i + \varepsilon_i \quad (2.2)$$

Table 2.7 presents estimates of Equation (2.2). Column 1 excludes covariates  $X_i$ , while column 2 adds them. The estimates are statistically identical, and the point estimates are very similar. For customers who received the standard coupon and were in the information control group, the CFL market share is 34 percent. The rebate increases CFL market share by about ten percentage points. This implies a price elasticity of demand for CFLs of  $\frac{\Delta Q/Q}{\Delta P/P} \approx \frac{0.1/0.34}{-0.2} \approx -1.5$ . Column 3 shows that the interaction between information and rebates is statistically zero.

The informational intervention does not statistically affect CFL market share. Using the standard errors from column 2, we can reject with 90 percent confidence that the intervention had more than 73 percent of the effect of the 20 percent CFL rebate. Assuming

---

<sup>12</sup>We technically prefer the linear probability model here because we assume locally linear demand when using the estimates for policy analysis. In any event,  $S$  and  $T$  are indicator variables, and the probit estimates are almost identical.

**Table 2.7:** *Effects of In-Store Informational Intervention*

	(1)	(2)	(3)
1(Rebate)	0.094 (0.035)***	0.105 (0.033)***	0.078 (0.047)*
1(Treatment)	-0.002 (0.035)	0.004 (0.033)	-0.022 (0.045)
1(Rebate and Treatment)			0.054 (0.066)
R2	0.01	0.16	0.16
N	794	793	793
Individual Characteristics	No	Yes	Yes

*Notes:* This table presents estimates of Equation (2.2), a linear probability model with outcome variable 1(Purchased CFL). Robust standard errors in parenthesis. \*, \*\*, \*\*\*: Statistically significant with 90, 95, and 99 percent confidence, respectively.

linear demand, this bounds the effect of information at the effect of a 12 percent rebate, or about \$0.46 per 60-Watt equivalent bulb.

There are several reasons for why the information effect might differ from the TESS experiment. First, it could be that a very large share of consumers did not understand the in-store informational intervention or were in too much of a hurry to internalize the information. However, our RAs reported that they believe that this is unlikely. Second, the informational environment differs: these and other home improvement stores have signage in lightbulb aisles that highlights features of different lightbulb technologies, including electricity use. If this existing information is very effective, incremental information might have no effect. Notice that if this is the case, our treatment effects are still the relevant parameters for policy analysis later in the paper: if existing information provision technologies are fully effective, then there is no remaining imperfect information and inattention to justify subsidies and standards.

A third reason is that the experimental populations differ: the TESS population is nationwide, while the in-store sample is drawn from four stores in three eastern states. Home improvement retailers are the most common place where households buy lightbulbs

(DOE 2010), and our partner alone sells upwards of 50 million lightbulb packages each year, a non-trivial share of national sales. Internally valid estimates for our experimental sample are thus of great interest *per se*, even if the results might not generalize to other types of retailers.

## 2.5 A Framework for Policy Analysis

### 2.5.1 Consumers

We model consumers that make one of two choices, labeled  $E$  and  $I$ . In our empirical application,  $E$  represents the purchase of an energy efficient product (the CFL), while  $I$  is an energy inefficient product (the incandescent). More generally, this model could capture any choice over which consumers might misoptimize.

Products  $j \in \{E, I\}$  are sold at prices  $p_j$ , and  $p = p_E - p_I$  is the relative price of  $E$ . We define  $v_j$  as the consumer's true utility from consuming product  $j$ , and we call  $v = v_E - v_I$  the relative true utility from  $E$ . In our empirical application,  $v$  could be determined by any and all of the differences between CFLs and incandescents, such as electricity costs, longer lifetime, mercury content, brightness, and "warm glow" utility from reduced environmental impact.

A consumer's utility from purchasing product  $j$  at price  $p_j$  is given by  $v_j + (Z - p_j)$ , where  $Z$  is the consumer's budget and  $Z - p_j$  is utility from consumption of the numeraire good. A fully optimizing consumer thus chooses  $E$  if and only if  $v > p$ . A misoptimizing consumer chooses  $E$  if and only if  $v - b_k(p) > p$ , where  $b_k(p)$  is a bias that may depend on  $p$  and is continuously differentiable in  $p$ . To simplify notation, we will typically denote bias by  $b_k$  rather than  $b_k(p)$ . The pair  $(v, b_k)$ , which we will refer to as a consumer's "type," is jointly distributed according to a distribution  $F$ . We denote the conditional distribution of  $v$  given  $b$  by  $F_{v|b_k}(\cdot|b_k)$ . We assume that for each  $b_k$ , the distribution  $F_{v|b_k}(\cdot|b_k)$  has an atomless and continuous density function  $f_{v|b_k}(\cdot|b_k)$ . For simplicity, we will also assume that  $b$  takes on finitely many values, though the analysis easily generalizes. We let  $\alpha_k$  denote the fraction

of consumers with bias  $b_k$ .

We also call  $b_k$  the "internality," to highlight the analogy to externalities. While an externality is a wedge between private willingness-to-pay (WTP) and social welfare, the internality is a wedge between private WTP and true private welfare. This is a reduced form representation of many biases that could cause consumers not to maximize experienced utility, including misperceptions of any product attribute. It allows for dependencies between bias  $b_k$ , true valuation  $v$ , and price  $p$ , as theories of endogenous inattention would imply. In our empirical application, we think of the bias as arising from consumers' undervaluation of energy costs due to a set of informational and attentional biases that we discuss in the next section.

Under the additional assumption below, this model generates continuous and downward-sloping demand curves for product  $E$ .<sup>13</sup>

**Assumption 1:**  $b_k$  is differentiable in  $p$  and there exists a  $\rho > -1$  such that  $b'_k(p) > \rho$  for all  $p$ .

Let  $D^R(p) = 1 - F(p)$  denote the "unbiased" demand curve for  $E$ , let  $D_{b_k}(p) = \alpha_k[1 - F_{v|b_k}(p + b_k|b_k)]$  denote the demand curve of consumers with bias  $b_k$ , let  $D_{b_k}^R = \alpha_k[1 - F_{v|b_k}(p|b_k)]$  denote what would be the demand curve of consumers with bias  $b$  if they were "debiased," and let  $D(p) = \sum_k D_{b_k}(p)$  denote the total demand curve of all consumers. Our assumptions about  $b_k$  and  $F$  imply that all demand curves are continuously differentiable functions of  $p$ . All analysis that follows expresses results in terms of these demand curves, so this framework could also be applied to continuous choice situations.

## 2.5.2 The Policymaker

The policymaker has two types of tax policies available: a subsidy of amount  $s$  for  $E$  and a ban on either choice. We will compare the welfare impacts of these policies to a hypothetical

---

<sup>13</sup>The assumption that  $b'_k(p) > -1$  is needed to guarantee that a consumer's perceived relative value of  $E$ , given by  $v - b_k(p) - p$ , does not increase in  $p$ . The additional assumption that  $b'_k$  is bounded away from -1 guarantees that  $v - b_k(p) - p < 0$  for high enough  $p$  and that  $v - b_k(p) - p > 0$  for low enough  $p$ , which implies that type  $k$  is marginal at some price.

technology that can fully debias consumers.

The policymaker maintains a balanced budget through lump-sum tax or transfer  $T(s) = \int s\sigma(v, b_k, c - s)dF(v, b_k)$ . This implies that the subsidy has no distortionary effects on other dimensions of consumption, and thus its role is purely corrective. Because all consumers choose either  $E$  or  $I$ , the subsidy for  $E$  is equivalent to a tax on  $I$ , and a ban on one choice is equivalent to a mandate for the other. Products  $E$  and  $I$  are produced in a competitive economy at a constant marginal costs  $c_j$ , with relative cost  $c = c_E - c_I$ . Product  $E$ 's relative price after subsidy  $s$  is  $p = c - s$ .

Let  $\sigma(v, b_k, p)$  denote the choice of a type  $(v, b_k)$  consumer at price  $p$ , with  $\sigma = 1$  denoting the choice of  $E$  and  $\sigma = 0$  denoting the choice of  $I$ . Also denote  $p_\sigma$  as  $p_E$  if  $\sigma = 1$  and  $p_I$  if  $\sigma = 0$ . Finally, define a normalizing constant  $C$  as the integral over all consumers of  $Z + v_I - p_I$ . The policymaker's objective is

$$\begin{aligned} W(s) &\equiv \int \overbrace{[\sigma(v, b_k, c - s)(v)]}^{\text{relative true utility from } E} \overbrace{[+Z + T(s) - p_\sigma]}^{\text{consumption of numeraire good}} dF(v, b_k) \\ &= C + \int \sigma(v, b_k, c - s)(v - c)dF(v, b_k). \end{aligned} \quad (2.3)$$

Our assumptions about  $b_k$  and  $F$  imply that  $W$  is continuously differentiable. To ensure the existence of an optimal subsidy, we assume that set of possible values of  $v - b_k$  is bounded from above and from below; that is,  $\bigcup_k \text{supp} F_{v|b_k}(\cdot|b_k)$  is a bounded set. This assumption implies that a ban on product  $I$  is equivalent to a sufficiently large subsidy, and a ban on product  $E$  is equivalent to a sufficiently small (negative) subsidy (i.e., sufficiently large tax on  $E$ ).

An important benchmark we will consider is the *first best level of welfare*, given by

$$W^{FB} = C + \int_{v \geq c} (v - c)dF. \quad (2.4)$$

## Subsidy

Define the *average marginal internality* at price  $p = c - s$  as

$$B(p) = \frac{\sum_k b_k D'_{b_k}}{D'}.$$

In Appendix B.4, we establish the following result, which is analogous to the optimal tax formulas derived by Allcott, Mullainathan and Taubinsky (2013) in a similar setting:

### Proposition 15.

$$W'(s) = (B(c - s) - s)D'(c - s) \quad (2.5)$$

The intuition behind Proposition 15 is that the welfare impact of a subsidy trades off the internality reduction,  $B(c - s)D'$ , with the distortion to consumers' decision utility,  $-sD'$ . This is directly analogous to the logic behind a Pigouvian tax, which trades off externality reduction with distortions to consumers' private utility gains. This result is closely related to the formulas derived in Allcott, Mullainathan and Taubinsky (2013), Baicker, Mullainathan, and Schwartzstein (2013), and Mullainathan, Schwartzstein, and Congdon (2012) for similar settings.

Since an optimal subsidy  $s^*$  must satisfy  $W'(s^*) = 0$ , Proposition 15 implies that an optimal subsidy must equal the average marginal internality:

$$s^* = B(c - s^*) \quad (2.6)$$

Equation (2.6) is analogous to Diamond's (1973) result that the optimal externality tax when agents have heterogeneous externalities is a similarly-weighted average marginal externality.

Figure 2.3 illustrates the intuition in the special case with linear demand and constant  $B(p)$ . The unbiased demand curve  $D^R(p)$  is shifted out relative to demand curve  $D(p)$ , meaning that the bias reduces demand for good  $E$  and causes welfare losses. The average marginal internality is the distance from  $a$  to  $f$ . The initial equilibrium is at point  $b$ , and the optimal subsidy  $s^*$  moves the equilibrium to point  $f$ . A marginal increase in the subsidy from 0 induces marginal consumers at point  $b$  to purchase good  $E$ , increasing their true utility by amount  $bd$ . The welfare gain from the optimal subsidy is triangle  $abd$ .

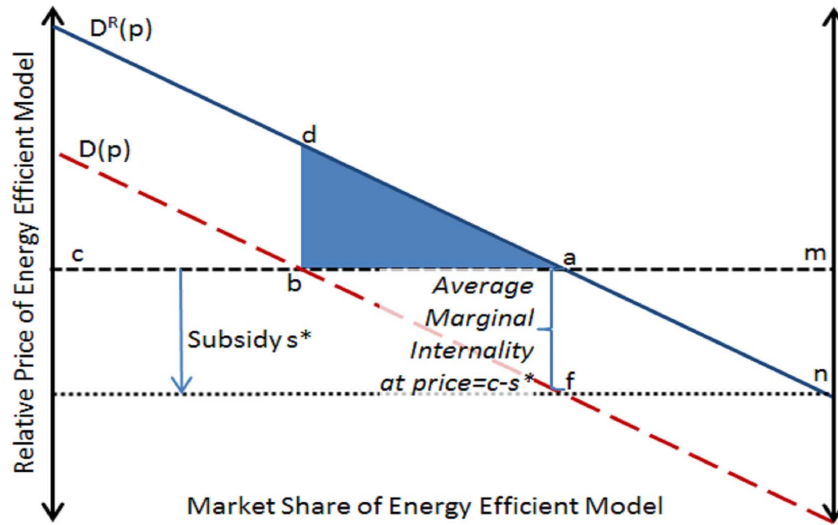


Figure 2.3: Marginal Internality Example

Equation (2.6) highlights the kinds of consumer heterogeneity that do and do not matter for the optimal subsidy. Heterogeneity in the average marginal internality  $B(p)$  at different price levels clearly does matter, and there are a number of practical situations where one might expect higher-WTP consumers to be more or less biased. For example, environmentalist consumers may both be more attentive to energy costs and have higher true relative utility  $v$  from good  $E$ . This highlights that it is not sufficient to set an internality tax based on a general idea that "consumers are biased" - it matters whether the biased consumers are marginal to the policy. As the analogy to Diamond's (1973) formula suggests, this insight generalizes to other market failures: if setting a time-invariant congestion tax, for example, the optimal tax would be smaller if people who travel at rush hour (and thus impose larger externalities) are less price elastic.

However, heterogeneity in bias  $b$  within the set of consumers on the margin at price  $p$  does not affect the optimal subsidy: only the average marginal internality matters. This is important because some models such as Chetty, Looney, and Kroft (2007) have consumers that are either fully unbiased or completely biased with some probability, while other models might have all consumers with a partial bias. While empirical analyses such as



Chetty, Looney, and Kroft (2009), Hossein and Morgan (2006), and Abaluck and Gruber (2011) have been able to identify average biases within groups of marginal consumers, we are not aware of previous studies that have been able to identify distributions of individual biases. Equation (2.6) shows that the optimal subsidy can be set without knowledge of the underlying "structural" model and distribution of biases. Thus, the  $B(p)$  function is a sufficient statistic for setting the optimal subsidy in the sense of Chetty (2009).

While the  $B(p)$  function is a sufficient statistic for calculating the welfare impacts of a subsidy, it is not informative about how close the optimal subsidy comes the first best. As the next proposition shows, a subsidy can achieve the first best if and only if consumers have homogeneous bias. Combined with the fact that the same  $B(p)$  function can be generated by either homogeneously or heterogeneously biased consumers, this proposition implies that  $B(p)$  is not informative of the gap between the second best welfare attained by the optimal subsidy and the first best level of welfare that would be attained if consumers were fully debiased.

**Proposition 16.** *Let  $s^*$  be an optimal subsidy. If all consumers have the same bias  $b$ , then  $s^*$  is uniquely defined and  $W(s^*) = W^{FB}$ . If some consumers have bias  $b_i$  while other consumers have bias  $b_j$  such that  $b_i(p) < b_j(p)$  for all  $p$ , then  $W(s^*) < W^{FB}$ .*

A key implication of Proposition 16 is that if an informational intervention that fully debiases consumers were inexpensive and feasible at large scale, it would likely be preferred to a subsidy or ban. If bias  $b$  is heterogeneous, a subsidy is less efficient: if it perfectly corrects the choice of consumers with bias  $b_1$ , it still leaves consumers with bias  $b_2 > b_1$  to underpurchase  $E$ , while leaving consumers with bias  $b_3 < b_1$  to overpurchase  $E$ . This is the intuition for why "asymmetric paternalism" (Camerer *et al.* 2003) and "libertarian paternalism" (Sunstein and Thaler 2003) are preferred to subsidies and bans. The reason to also consider subsidies and bans is if fully-debiasing information provision technologies are costly or infeasible, while the feasible information provision technologies do not fully debias consumers. In particular, our informational interventions are unlikely to be scaled: even though we chose high-volume store locations, the in-store experiment required labor costs

of several dollars per customer intercept, and it is not obvious how the TESS intervention could be implemented outside the TESS platform.

## Ban

According to welfare equation (2.3), a ban on good  $I$  has the following effect on welfare:

$$\begin{aligned}
 \Delta W &= \int (v - c) dF(v, b_k) - \int \sigma(v, b_k, c) (v - c) dF(v, b_k) \\
 &= \int (1 - \sigma(v, b_k, c)) (v - c) dF(v, b_k) \\
 &= \int_{\{(v, b_k) | \sigma(v, b_k, c) = 0\}} v dF(v, b_k) - c \tag{2.7}
 \end{aligned}$$

This equation simply states that the welfare effects of a ban are the average relative true utility  $v$  for consumers currently purchasing  $I$ , minus the relative cost  $c$ . An analogous equation would hold for a ban on  $E$ .

Figure 2.3 illustrates this equation, again assuming linear demand and constant  $B(p)$ . The ban on good  $I$  increases welfare for the set of consumers to the left of point  $f$ , because purchasing good  $E$  increases their true utility. This welfare gain is the triangle  $abd$ . However, the ban decreases welfare for the set of consumers to the right of point  $f$ : while they are biased, their true utility from good  $E$  is still less than the relative price. This welfare loss is the triangle  $amn$ .

Under our assumption of lump-sum revenue recycling, a ban on good  $I$  is equivalent to a sufficiently large subsidy for good  $E$ . Bans are thus weakly worse than the optimal subsidy. However, there is some marginal cost of public funds at which a ban could be preferred. On the other hand, if the corrective price policy is implemented as a tax on good  $I$ , then a cost of public funds further reinforces the relative appeal of price-based policies relative to bans.

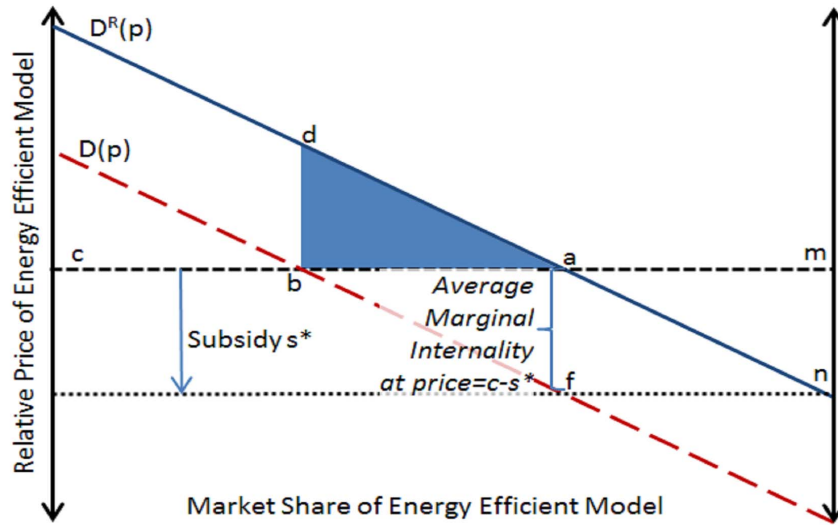


Figure 2.4: Demand Curves

### 2.5.3 First-Order Approximation to Optimal Subsidy

When the policymaker can directly measure the  $B(p)$  function at all  $p$ , welfare changes can be computed at each subsidy level to exactly compute the globally optimal subsidy. In Section 6, we use the TESS experiment to do this. However, this function is in general difficult to estimate: there are only a few papers in any context that cleanly identify biases for even some subset of consumers. Following Baicker, Mullainathan, and Schwartzstein (2013) and Mullainathan, Schwartzstein, and Congdon (2012), we now present one approach to approximating the marginal internality using two reduced form sufficient statistics: the slope of demand and the effect of the bias on market shares. In Section 6, we illustrate how this can be implemented using the in-store experiment.

To a first order approximation, we have

$$\begin{aligned}
 D_{b_k}^R(p) - D_b(p) &= D_{b_k}(p - b_k) - D_{b_k}(p) \\
 &\approx b_k D'_{b_k}(p).
 \end{aligned}$$

Thus,

$$D^R(p) - D(p) = \sum_k (D_{b_k}(p - b_k) - D_{b_k}(p)) \approx \sum_k b_k D'_{b_k} = D'(p)B(p).$$

It then follows that

$$B(p) \approx \frac{D^R(p) - D(p)}{D'(p)}. \quad (2.8)$$

The numerator is the effect of the bias on market shares, while the denominator is the slope of demand. In other words, the average marginal internality is the price change that affects quantity demanded exactly as much as the bias does.

To a first-order approximation, demand and "unbiased demand" have the same slope:  $(D^R_{b_k})'(p) = D'_{b_k}(p - b_k) \approx D'_{b_k}(p)$ . This means that we can also approximate  $B$  by

$$B(p) \approx \frac{D^R(p) - D(p)}{(D^R)'(p)}. \quad (2.9)$$

Figure 2.3 illustrates the intuition behind Equation (2.8). The length of segment  $ab$  is given by  $D^R(p) - D(p)$ . This could be identified through a randomized field experiment that fully debiases consumers and measures the effects on demand for  $E$ . The demand slope  $D'(p)$  could be identified through an RCT that randomizes relative prices. The segment  $af$ , which corresponds to the average marginal internality, is given by  $\frac{D^R(p) - D(p)}{D'(p)}$ . Combining Equation (2.8) or (2.9) with Equation (2.6) yields an approximation to the optimal subsidy. Thus, the effect of the bias on market shares and the slope of demand are sufficient statistics for an approximation to the average marginal internality.

## 2.6 Policy Evaluation

### 2.6.1 Inferring Bias from Treatment Effects

We now combine the experimental estimates with the theoretical framework to illustrate how results such as these could be used to inform policy. To do this, we build on the idea that the information treatment groups choose optimally, although the information control group may not. One important feature of this approach is that it "respects choice" in the

sense of Bernheim and Rangel (2009): we conduct welfare analysis using consumers' own choices in what is plausibly a "debiased" state. In their language, we define control group choices as *provisionally suspect* due to the possibility of imperfect information processing. If choices differ between treatment and control, we delete control group choices from the welfare-relevant domain. In our language, the implication is that the conditional average treatment effect of our informational interventions at any price  $p$  equals the average marginal internality from imperfect information and inattention:

**Assumption 2:**  $\tau(p) = B(p)$ .

This is analogous to the assumption made by Chetty, Looney, and Kroft (2009) when they estimate the magnitude of inattention to sales taxes using the treatment effect of an intervention that posted tax-inclusive purchase prices. In justifying this assumption, Chetty, Looney, and Kroft (2009) write that "when tax-inclusive prices are posted, consumers presumably optimize relative to the tax-inclusive price." Similarly, it seems reasonable to assume that consumers optimize relative to lightbulb lifetimes and energy costs after we provided them with information about these attributes. In the empirical sections, however, we have discussed potential reasons why this assumption may not hold. In qualitatively interpreting the results, we view this assumption as an approximation.

## 2.6.2 "Structural" Models of Bias

Because the optimal policy depends on  $b$ , not the underlying "structural" model of the bias, our exposition uses this "reduced form" parameter. However, it may be helpful to specify categories of inefficiencies that could affect lightbulb demand and would plausibly be addressed by our informational interventions:

1. Costly information acquisition, as in Gabaix *et al.* (2006) and Sallee (2013). This category includes many standard models of imperfect information in which the consumer incurs a cost to learn about energy efficiency, lifetime, or other product attributes and, in the absence of paying that cost, assumes that different goods have the same level of an attribute.

2. Biased priors about energy costs or other product attributes, as tested by Allcott (2013), Attari *et al.* (2010), Bollinger, Leslie, and Sorensen (2011), and others. Put simply, this category reflects consumers who may know that CFLs use less energy but don't know that the savings are so large.
3. Exogenous inattention to energy as a "shrouded" add-on cost, as in Gabaix and Laibson (2006).
4. Costly cognition or "thinking cost" models, as in Conlisk (1988), Chetty, Looney, and Kroft (2007), Gabaix (2013), Reis (2006), Sims (2004), and others. In these models, consumers might not pay attention to differences in energy costs between lightbulbs because their experiences with other goods suggest that energy cost differences are typically unimportant. However, once informed that lightbulb energy cost differences are large relative to the difference in purchase prices, consumers in these models would consider them in their choices.

Informational interventions would not affect all biases that could affect lightbulb demand. For example, "bias toward concentration" (Koszegi and Szeidl 2013) could cause consumers to undervalue electricity costs because they occur in a stream of small future payments. Koszegi and Szeidl (2013) point out that re-framing the stream of payments as one net present value, as our interventions do, does not necessarily address this possible bias. It is also possible that present bias over cash flows could cause consumers to undervalue the CFL's future cost savings, although this is not consistent with the TESS data or the standard models of present bias over consumption. Our informational interventions should not affect present biased consumers. Finally, consumers could be imperfectly informed about or inattentive to other attributes not discussed in our informational interventions.

Because imperfect information and inattention may not be the only biases, Appendix B.4 generalizes the theoretical framework to the case when the intervention identifies only part of the bias. Intuitively, the optimal subsidy is additive in the different types of internality. For example, if present bias over cash flows reduces marginal consumers' demand for

the CFL, then the true optimal subsidy is larger than the subsidies we calculate based on the informational interventions alone. Similarly, if present bias reduces inframarginal consumers' CFL demand, then the true welfare gains of a ban are larger (or less negative) than we calculate.

### 2.6.3 Using the TESS Experiment Results

#### Subsidy

The theoretical framework shows that to set and evaluate policy, we need to know both the initial demand curve and the average marginal internality at each point. This should now clarify why the particular design we used for the TESS experiment is so important: by eliciting WTP in consumers' baseline (potentially biased) state and subsequently in their treated (optimizing) state, we can identify the average marginal internality at each point on the market demand curve.

Figure 2.4 presents the conditional average treatment effects (CATEs) at each level of baseline WTP. As Figure 2.1 shows, there are only a small number of consumers with baseline WTP equal to \$9 or between -\$3.50 and -\$9, so we group outlying high and low baseline WTPs together. Consistent with the ATEs in Table 2, the CATEs are all around \$3, except for the CATE at the highest baseline WTP, which is close to zero. This is simply due to top-coding: consumers who start with top-coded baseline WTP cannot increase their WTP further. Because these inframarginal consumers are unaffected by the subsidy and the ban, this does not affect the welfare calculations. After excluding consumers with top-coded and bottom-coded baseline WTP, there is a slight positive correlation between baseline WTP and the treatment effect. This highlights that the population average internality would not be the right parameter for setting optimal policy.

Figure 2.5 illustrates how this distribution of average marginal internalities is combined with the baseline demand curve for policy analysis. The dashed line is the baseline demand curve  $D(p)$ . At each point, the average marginal internality from Figure 2.4 is added to WTP

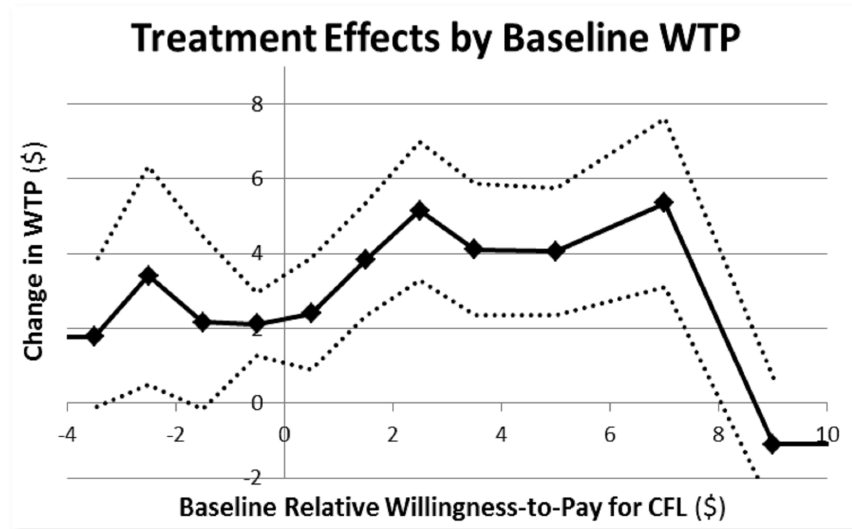


Figure 2.5: TESS Experiment Treatment Effects by Initial WTP

to give the average true utility of consumers marginal at each price  $p$ . These average true utilities are plotted as diamonds. Consistent with their approximately equal retail prices, we assume that the two packages have the same marginal cost of production, so  $c = 0$ .

The leftmost shaded rectangle reflects the welfare gain from increasing the subsidy from \$0 to \$1. This increased subsidy moves about 13 percent of consumers over the margin to buying a CFL. The height of the rectangle is the difference between average true utility and relative cost  $c$ . Moving to the right, the next two shaded rectangles reflect the welfare gain from increasing the subsidy from \$1 to \$2 and from \$2 to \$3, respectively. These first three shaded rectangles reflect welfare gains, as average true utility  $v$  exceeds cost  $c$ . However, further increases in the subsidy cause welfare losses. The \$3 optimal subsidy is consistent with Figure 2.4, which shows ATEs in the range of \$2 to \$3 for consumers with baseline WTP less than 0.

The welfare effects of banning incandescents are the sum of all shaded rectangles. Notice again that it is the average marginal internalities that determine welfare impacts, not the distribution of individual biases within the sets of marginal consumers. Graphically, this is reflected by the fact that height of each discrete welfare rectangle is determined only by the



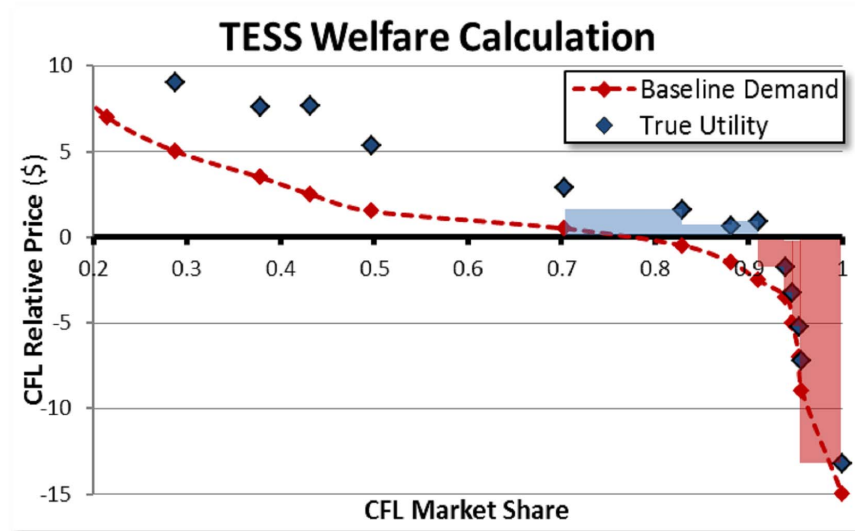


Figure 2.6: Welfare Calculation Using TESS Experiment

average internality at each price.

Table 2.8 presents formal calculations that parallel Figure 2.5. Column 1 contains the subsidy amount. Column 2 presents the average baseline relative WTP  $v^0$  for consumers marginal to the increase in the subsidy, assuming that demand is linear between the two price levels. Column 3 presents the average internality for this group of marginal consumers. This equals the treatment effect for each point on the left side of Figure 2.4. Column 4 presents the demand density: the share of all consumers that are marginal at each relative price level. Column 5 presents the welfare effect of the increment to the subsidy, using Equation (2.5), while column 6 presents the total welfare effect of changing the subsidy from zero to the amount listed in that row. Columns 3-6 are measured with sampling error, although we omit standard errors for simplicity.

As Equation (2.5) shows, a marginal increase in the subsidy increases welfare as long as the marginal internality outweighs the distortion to decision utility. Table 2.8 shows that for subsidies larger than \$3, the point estimate of true utility for marginal consumers is less than the cost of the CFL. Thus, increases in the subsidy above \$3 will reduce welfare relative

**Table 2.8:** *Welfare Analysis Using TESS Results*

(1)	(2)	(3)	(4)	(5)	(6)
CFL Subsidy (\$/package)	Relative WTP ( $v_i^0$ ) of Marginal Consumers (\$/package)	Average Marginal Internality (\$/package)	Demand Density (Share of packages)	Marginal Welfare Effect (\$/package)	Cumulative Welfare Effect (\$/package)
1	0.5	2.11	0.126	0.204	0.204
2	1.5	2.16	0.052	0.034	0.238
3	2.5	3.41	0.028	0.026	0.264
4	3.5	1.77	0.030	-0.052	0.212
6	5	1.77	0.006	-0.020	0.192
8	7	1.77	0.008	-0.042	0.150
10	9	1.77	0.003	-0.019	0.131
$\infty$	15	1.77	0.043	-0.567	-0.436

*Notes:* This table uses the TESS experiment results to calculate the welfare effects at different levels of the CFL subsidy. Observations are weighted for national representativeness. See text for details.

to the \$3 subsidy in our model. Using the analogy to externalities, subsidies higher than \$3 would be equivalent to setting a Pigouvian externality tax higher than marginal damages. For comparison, typical CFL rebates offered by electric utilities have been on the order of \$1 to \$2 per bulb.

If we assume that  $B(p) = \hat{\tau}(p) = \hat{\tau} = \$2.32$  at all  $p$ , Equation (2.6) implies that the globally optimal subsidy is also  $s^* = \$2.32$ , which is consistent with the result in Table 2.8 that the optimal subsidy does not exceed \$3. However, the benefit of this "grid search" approach to calculating the optimal subsidy is that in theory, the average marginal internality could be very different for different values of  $v^0$ . Grid search identifies the global optimum even if the necessary condition for a local optimum in Equation (2.6) is satisfied at multiple subsidy levels.

## Ban

A ban on traditional incandescents is equivalent to a change in relative prices that is so large as to induce all consumers to cease buying incandescents. In Figure 2.5, this is the sum of the positive welfare rectangles above the x-axis minus the sum of the negative welfare

rectangles below the x-axis. Table 2.8 shows that this sums to a loss of \$0.436 per package sold.

Top-coding and bottom-coding have two opposing effects on this welfare calculation. First, the treatment causes many treatment group consumers to be willing to pay the maximum for the CFL. Assuming a larger average WTP for this top-coded group would increase the treatment effect, implying a larger bias and thus larger welfare gains from corrective policies. Second, however, the welfare effects of a ban depend importantly on the tail of the WTP distribution: if some consumers very strongly prefer incandescents, banning them can cause large welfare losses. Appendix Table B2.3 tests the sensitivity of these welfare calculations to assuming that top-coded and bottom-coded relative WTPs average \$12 and -\$12. The two opposing effects almost exactly offset each other: the welfare loss from the ban is \$0.434 per package sold. Appendix Table B2.3 instead assumes mean censored WTPs of \$20 and -\$20. Under this assumption, the welfare loss from the ban is \$0.744 per package sold.

### **Illustrative Calculation: Welfare Gains from Information Provision**

Proposition 16 shows that when the bias is heterogeneous, a fully-debiasing informational intervention generates larger consumer welfare gains than a subsidy. For the purpose of illustrating this, we briefly make a very strong assumption: that each treated consumer's individual WTP change from baseline to endline equals  $b$ . This is stronger than our Assumption 2, which was that the conditional average treatment effects equal  $B(p)$ .

Debiasing changes a consumer's welfare if and only if it changes his or her choice at market prices. The utility gain for a consumer who does change his or her choice is  $|v - c|$ . Thus, the welfare gain from full debiasing is the integral of  $|v - c|$  over all consumers who change choices:

$$\int |\sigma(v, 0, c) - \sigma(v, b_k, c)| \cdot |v - c| dF(v, b_k). \quad (2.10)$$

In the TESS data, 21 percent of treatment group consumers change choices after the

intervention, and their average  $|v - c|$  is \$3.36, giving total welfare gains of \$0.72 per package.<sup>14</sup> This is almost three times larger than the \$0.26 per package welfare gain from the optimal subsidy. Of course, to fully compare the two policies, one would need to subtract the cost of implementing each policy, including the costs of consumers' time for the informational intervention and any deadweight loss of raising public funds for the subsidy. This calculation simply illustrates the sense in which the "targeting" properties of information provision can make it preferred to a subsidy.

## 2.6.4 Using the In-Store Experiment Results

The treatment effects from the in-store experiment can be used in Equation (2.8) to determine the optimal subsidy for this sample in this context. If the treatment group optimizes with respect to energy costs and product lifetimes, while the control group is potentially biased,  $\hat{\tau} = D^R(p) - D(p)$ . We assume that  $D' = (D^R)'$ , as we do not reject this hypothesis in Table 2.7, and it is theoretically true to first-order approximation. Plugging  $\hat{\tau}$  and  $\hat{\eta}$  from column 2 of Table 2.7 into Equation (2.8) and using that the average rebate  $s$  per 60-Watt bulb was \$0.63, the optimal subsidy per 60-Watt bulb is:

$$B(S = 0) \approx \frac{D^R(p) - D(p)}{D'(p)} = \frac{\hat{\tau}}{\hat{\eta}/s} \approx \frac{0.004}{0.105/\$0.63} \approx \$0.024 \quad (2.11)$$

Figure 2.6 illustrates the calculation. The information treatment effect  $\hat{\tau}$  is the distance from  $b$  to  $a$ , and the slope of demand is  $\hat{\eta}/s$ . The figure exaggerates  $\hat{\tau}$ , as the point estimate suggests only a very small effect on market share.

Using the Delta method and the estimated variance-covariance matrix, the 90 percent confidence interval on the optimal subsidy per 60-Watt equivalent CFL is  $(-\$0.30, \$0.35)$ . Thus, for this sample of people in the informational environment where our experiment

---

<sup>14</sup>We emphasize that this calculation is purely for the purposes of illustrating Proposition 16. Figure 2.2 illustrates why the required assumption is too strong: some control group consumers also change WTP between baseline and endline, even though the control intervention was not designed to debias. Under a similar assumption, Equation (2.10) would imply that the welfare gains from the control intervention are \$0.19. This suggests that even if the conditional average treatment effects are meaningful, there can be noise in any given consumer's WTP change between baseline and endline.

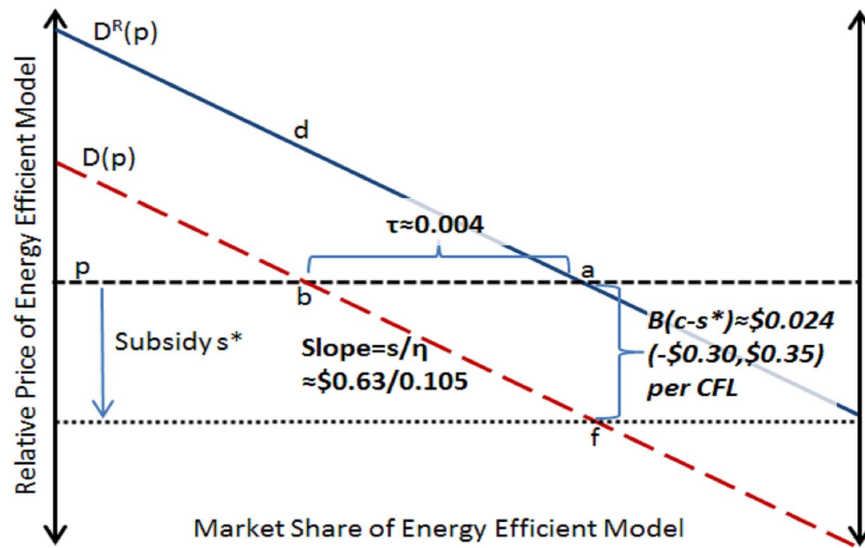


Figure 2.7: Policy Analysis Using In-Store Experiment

took place, the results rule out that the optimal subsidy is more than one-third as large as the \$1 to \$2 per bulb subsidies that electric utilities have typically offered. Furthermore, unless the treatment group consumers who bought incandescents in this experiment have substantially weaker preferences for incandescents than the inframarginal consumers in the TESS experiment, the tightly-estimated zero treatment effect suggests that banning incandescent lightbulbs will cause larger welfare losses in this population than in the TESS population. Thus, while the empirical estimates from the two experiments are different, they both lead to the same qualitative conclusion about the ban.

## 2.7 Conclusion

Imperfect information and inattention are commonly-proposed justifications for energy efficiency subsidies and standards, and they are frequently invoked in the lightbulb market as justifications for energy efficiency policies. We implemented two randomized control trials that measure the effects of "powerful" information provision on purchases of energy efficient

CFLs. The TESS intervention increased WTP by an average of \$2.32, while the in-store experiment had tightly-estimated zero effects in a different population and informational environment. These forms of information provision would be difficult to scale, and lower-cost disclosure technologies seem less likely to fully inform and debias consumers. We thus formalized a theoretical framework that uses the experimental results to evaluate two second best policies: CFL subsidies and a ban on traditional incandescent bulbs. Results suggest that moderate CFL subsidies may be optimal, but that imperfect information and inattention do not justify a ban on traditional incandescents.

For our quantitative evaluation of subsidies and standards, we assumed that the treatment effects of powerful information provision identify the magnitude of bias. Given our experimental designs, we think that the assumption is a reasonable but imperfect approximation. We have discussed the importance of issues such as whether the treatment groups understood the information, experimenter demand effects, and external validity.

Even though the policy analysis is approximate, there are several ways in which this study is valuable. First, the in-store experiment is a proof-of-concept for what we think will be an important research effort to use field experiments to evaluate the effects of energy efficiency subsidies and information provision on purchases of durable goods. Second, we have highlighted the necessary parameters for studying the "internality rationale" for energy efficiency policies, and we have implemented two examples of experimental designs that can identify these parameters. Third, as we calculated in Section 2, estimates of inattention from other research combined with the very large magnitude of lightbulb energy costs relative to purchase prices suggested that biases from imperfect information and inattention might have had large effects in this market. Results from both experiments are consistent in that they reject that potential prior. Fourth, our model provides one initial data point which suggests that imperfect information and inattention may not justify the lighting energy efficiency standards in the absence of other distortions.

Although our application is to one particularly important and controversial policy, the approach is quite general. The theoretical framework generalizes immediately to any binary

or continuous choice, and the idea of using informational interventions in RCTs to quantify internalities can be used in a wide variety of contexts. The approach to optimal policy could be useful in other contexts where informational or attentional biases might be used to justify subsidies or bans and where powerful information provision is feasible in small RCTs but not cost effective at large scale.

## Chapter 3

# Market Experience is a Reference Point in Judgments of Fairness<sup>1</sup>

### 3.1 Introduction

Empirical evidence shows that people's aversion toward unfair transactions can play an important role in markets and negotiations. In product markets, consumers' feelings of entitlement restrict sellers' ability to exploit changes in supply and demand (Kahneman et al., 1986), while in labor markets, reciprocal gift exchange can lead to involuntary unemployment (Akerlof, 1982; Fehr et al., 1993).<sup>2</sup> To incorporate such non-pecuniary concerns into economic theory, economists have proposed models of "social preferences," which assume that in addition to maximizing consumption, people also care about the fairness or kindness of own or others' actions. A common property of these models is that the fairness or kindness of an action or outcome is evaluated by an exogenous and static criterion such as equal division or surplus maximization, which implies that fairness judgments remain stable over

---

<sup>1</sup>Co-authored with Holger Herz

<sup>2</sup>For reviews of this evidence, see Fehr et al. (2009) on labor markets and Camerer (2003) on negotiations.



time and past experiences should not affect the evaluation criterion.<sup>3</sup>

Casual observation and introspection, however, suggest that this static description of people's feelings of entitlement is incomplete. Consider two examples. Americans are getting increasingly upset about raises in gasoline prices.<sup>4</sup> The Swiss have signed petitions to start referenda over a CHF 40,000 (roughly 41,500 USD) minimum yearly wage and an initiative that restricts the highest salary in a Swiss company to no more than 12 times the lowest one.<sup>5</sup> It appears that Americans feel entitled to cheap petrol, and the Swiss are accustomed to high wages. For outsiders, these expressions of anger are hard to comprehend. After all, European gas prices are twice as high as American prices, and the Swiss workers are already amongst the most privileged in the world in terms of wages. The anger, therefore, seems to be best explained by fairness standards that are dynamic and shaped by past experience.

Yet to our knowledge, direct evidence for *path-dependent fairness standards* is still missing. In this paper, we provide an experimental test of path-dependent fairness preferences by exogenously manipulating subjects' experiences, *keeping everything else constant*. To derive our hypotheses, we introduce a model of path-dependent fairness preferences in which past experience shapes people's feelings of entitlement. Our investigation is inspired by the seminal work of Kahneman et al. (1986), who introduce an intuitive notion of fairness in which "A firm is not allowed to increase its profits by arbitrarily violating the entitlement of its transactors to the reference price, rent, or wage," and who argue that past experience determines this reference transaction: "When there is a history of transactions between

---

<sup>3</sup>See, for example, Rabin (1993), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Dufwenberg and Kirchsteiger (2004), or Falk and Fischbacher (2006). Models including reciprocity motives suggest a certain context dependence since the desire to treat someone kindly depends on how they acted. But the evaluation of an agent's kindness still requires a static criterion. Different from these commonly used models of social preferences, but similar to ours, Benjamin (2005) proposes a model of reference-dependent fairness preferences in which an employee's period  $t$  wage serves as his period  $t + 1$  reference point, and Kaur (2012) formalizes the idea that workers may retaliate against a firm that offers them a wage below their reference wage.

<sup>4</sup>Tom Geoghegan. 2011. "Why are Americans so angry about petrol prices?" BBC News. May 11, 2011, accessed July 9th, 2013, <http://www.bbc.co.uk/news/world-us-canada-13338754>.

<sup>5</sup>Tim Judah. 2013. "Angry Swiss aren't done slimming the fat cats" Bloomberg. March 4, 2013, accessed July 9th, 2013, <http://www.bloomberg.com/news/2013-03-04/angry-swiss-aren-t-done-slimming-the-fat-cats.html>.

firm and transactor, the most recent price, wage, or rent will be adopted for reference...” and “terms of exchange that are initially seen as unfair may in time acquire the status of a reference transaction.”<sup>6,7</sup>

In the first phase of our experiment, all subjects participate in one of two market games. In the *proposer competition* (PC) game (Roth et al., 1991), two *proposers* make an offer of a monetary allocation to one responder, who can choose to accept either one or zero of those offers. In the *responder competition* (RC) game (Fischbacher et al., 2009), one proposer makes an offer to two responders, who simultaneously choose whether or not to accept the offer, with one responder randomly selected to transact in the case that both offers are accepted. Consistent with previous evidence, market conditions have a large impact on the offers: in the PC game, competitive pressures force proposers to give up most of their surplus, while in the RC game, proposers keep most of their surplus.

In the second phase of the experiment, proposers and responders are matched one-on-one in a variant of the ultimatum game (Güth et al., 1982), and proposers again make offers to responders. Consistent with previous studies, responders are willing to reject an offer and forgo significant monetary gains to punish proposers making unfair offers. However, we find that responders’ experiences from the first part of the experiment are an important reference point for the types of offers they are willing to accept: In period 1 of the ultimatum game, the lowest acceptable offer of a responder who started in the PC market is 36% higher than the lowest acceptable offer of a responder who started in the RC market. That is, responders who started out in markets in which competition forces proposers to make very

---

<sup>6</sup>In the Kahneman et al. study, participants rate the fairness of wage and price changes by firms in hypothetical scenarios. The majority of questions deal with differing justifications for similar wage or price changes and analyze how the motives underlying the price change affect the fairness rating. Only one of the analyzed questions (Question 2, p.730) actually examines the effect of differences in experience on the acceptability of the same offer, and they find that it is more acceptable for a firm to offer a low wage to a novel worker when market conditions have changed than offering the same wage to a current employee in which case the new wage constitutes a wage cut. However, note that in this case different earnings may also be justified based on differences in productivity.

<sup>7</sup>See also Binmore et al. (1991) for an early suggestion as well as survey evidence that exposure to different kinds of bargaining environments can shape subjects’ perceptions of what is fair. See Zwick and Mak (2012) for a recent review of the determinants of fairness perceptions in bargaining games.

favorable offers to the responders have a much higher standard for what constitutes a fair and acceptable offer. Moreover, this difference is persistent: over the course of 15 periods of repeated play, this difference dissipates by only about one-half of its period-one value.

In contrast, proposers' behavior is much less influenced by their phase 1 market experience, as they quickly learn what responders are willing to accept and adjust their offers to maximize profits. This result is consistent with a key theoretical prediction of our model: because past experience affects players' motives for resisting unfair treatment, it should affect the behavior of players with little bargaining power who are in a position to be treated unfairly, but it should not affect the behavior of players with greater bargaining power, because they are less likely to have their fairness reference point violated.

Our results have a number of economic implications. First, a key implication of our work is that fairness preferences are endogenous to market structure. For example, how a consumer perceives high prices can depend on whether this consumer is accustomed to high or low prices. This prediction is consistent with the differential perception of the price of gasoline discussed earlier.<sup>8</sup>

Second, dynamically adjusting fairness standards imply that consumer outrage following a price hike may be impermanent, and will subside as consumers adopt the new price as their reference transaction. Consider, for example, the outrage that followed a 60% price increase by movie rental company Netflix in July 2011. Starting with angry outcries in various social media channels,<sup>9</sup> this outrage quickly turned into a loss of 800,000 members and a stock price that plummeted from \$291 to \$75 over the course of just three and a half months.<sup>10</sup> Yet Netflix did not lower prices, and casual observation of the company two years later in 2013—a gain of 3 million new customers, a stock price above \$200, and no sign of discontent over unreasonably high prices—might suggest that consumers eventually

---

<sup>8</sup>In a similar vein, Simonsohn and Loewenstein (2006) show that movers to a new city arriving from more expensive cities rent pricier apartments than those arriving from cheaper cities.

<sup>9</sup><http://edition.cnn.com/2011/TECH/social.media/07/22/social.media.outrage.taylor/index.html>

<sup>10</sup>Nick Wingfield and Brian Stelter. 2011. "How Netflix lost 800,000 Members, and Good Will." New York Times. October 24, 2011, accessed June 20th, 2013, <http://www.nytimes.com/2011/10/25/technology/netflix-lost-800000-members-with-price-rise-and-split-plan.html>.

adopted the new prices as the reference transaction.

Third, our findings can help explain empirical observations in labor markets. It is a well established fact that current labor market conditions have little effect on incumbent workers' wages (Beaudry and DiNardo, 1991; Bewley, 1999; Kaur, 2012). Because workers' fairness reference points depend on past experience, lowering wages of existing workers is difficult, but hiring new workers at lower wages can be feasible. Moreover, our results also provide a rationale for the long lasting effects of initial labor market conditions and starting wages upon entering a firm (Oreopoulos et al., 2012). Firms take workers' current fairness reference point into account during future wage renegotiations, which causes persistent differences in offered wages based on initial starting conditions. Our results thus provide a potential psychological mechanism underlying wage rigidity and persistent wage differentials.<sup>11</sup>

Fourth, our work contributes to a recent literature on contracts as reference points by Hart and Moore (2008) (see also Hart, 2009; Hart and Holmstrom, 2010; Fehr et al., 2011b; Schmidt and Herweg, 2012), who argue that a contract between two parties functions as a reference point that these parties use to evaluate the fairness of their subsequent outcomes.<sup>12</sup> Our notion of reference-dependent fairness is complimentary, and applies more broadly to environments in which parties do not have the opportunity to write a contract prior to choosing actions. In fact, Hart and Moore (2008) discuss extensions of their model in which reference points other than contractual terms affect parties' feelings of entitlement. Our work, therefore, paves the way toward more integrated models of reference-dependent fairness.

Finally, our work contributes to the debate about the effects of policy interventions on market outcomes. Falk et al. (2006) experimentally show that experimenter-imposed minimum wage laws can cause spillover effects, raising wages even after the removal of

---

<sup>11</sup>See Kaur (2012) and reference therein for recent evidence and a potential source of high unemployment volatility. See Shimer (2005) and Hall (2005) for work relating wage rigidity to unemployment volatility. Eliaz and Spiegler (2012) propose a model of reference-dependent preferences in a search and matching model that generates both wage rigidity and unemployment volatility.

<sup>12</sup>See Fehr et al. (2011a) and Bartling and Schmidt (2012) on the effect of contracts as reference points on subsequent contract renegotiation.

the minimum wage law, a finding that is also predicted by our theory and consistent with our results. However, Falk et al. (2006) cannot pin down whether this spillover effect is directly due workers' past experiences: because the minimum wage was imposed by the experimenter, it may have served as a signal of the experimenter's preferences or beliefs about social norms.<sup>13</sup>

Our paper is also broadly related to experimental economics work on learning spillovers across games. The work on learning spillovers in strategic interactions (Grimm and Mengel, 2012; Bednar et al., 2012; Cason et al., 2011) has shown how beliefs about opponents' play can be influenced by observations of play in similar games. But while these papers demonstrate the importance of learning spillovers in coordination games through belief- or best-response bundling, our results demonstrate spillovers that operate at the level of preferences.

More generally, our results link the study of social preferences to the psychology and economics literature that argues that preferences are not exogenously given, but are reference-dependent and *constructed* from past experience, expectations, or the decision context (Lichtenstein and Slovic, 2006; Kahneman and Tversky, 2000; Simonson and Tversky, 1992). But while most existing work on reference-dependence studies how reference points affect the pricing of risk, the tradeoffs between consumption and effort, or the tradeoffs between different dimensions of consumption, our work shows the importance of reference-dependence in determining the tradeoff between consumption and fairness.

The rest of the paper proceeds as follows. In section 2 we describe the games used in our experiment, and the rest of our experimental design. In section 3 we present a simple and

---

<sup>13</sup>That is, experimenter-imposed wage conditions could have created a demand effect or led subjects to make inferences about social norms or what the experimenter considers to be appropriate behavior. In contrast, market forces endogenously shape experiences in our design, which allows us to directly attribute observed treatment differences to experiences. Moreover, because the Falk et al. (2006) design involves a complex strategic interaction between a firm and multiple workers, the design cannot pin down whether the outcomes of minimum wage laws shape fairness preferences or simply change players' beliefs about others' strategic intentions. (In their design, i) a firm can make up to three wage offers ii) payoffs are nonlinear in the number of workers accepting and iii) the multiple workers must simultaneously choose to accept or reject.) For example, how much a worker punishes a firm by rejecting its offer will depend on his belief about other workers' acceptance decisions, and his beliefs about the number of offers made by the firm. Or relatedly, workers who do not want to receive payoffs lower than other workers will not want to obtain zero payoffs by rejecting a wage offer if they think that other workers will accept offers and thus obtain positive payoffs.

generally applicable model of path-dependent fairness preferences, which we then use to motivate a set of hypotheses for our experimental design. In section 4 we present our results, and find that they are largely consistent with our theoretical predictions. In section 5 we discuss additional applications and testable predictions of our theory. Section 6 concludes.

## 3.2 Experimental Design

All games in the experiment were based on the asymmetric ultimatum game, first introduced by Kagel et al. (1996)<sup>14</sup>, and the market game first introduced by Roth et al. (1991). In each of these games, 100 chips must be divided between proposers and responders, with proposers making offers, and responders choosing whether or not to accept the offers. These chips are then converted into monetary payoffs, with different conversion rates for the proposer and the responder. In our experiment, the monetary value of each chip was three times as high for a proposer as it was for a responder.<sup>15</sup> Our experimental design consists of three variants of the asymmetric ultimatum game: (i) Proposer Competition (PC), (ii) Responder Competition (RC) and (iii) no competition. Subjects participated in one of the two market games for the first 15 periods of our experiment, and then participated in the non-competitive ultimatum game in the next 15 periods. We describe the experimental games in more detail below.

### 3.2.1 Phase 1: Market Games

In the first phase of our experiment (first 15 periods), subjects participated in either a responder competition treatment or in a proposer competition treatment.

In the responder competition (RC) market game, one proposer is matched with two

---

<sup>14</sup>This asymmetric ultimatum game is a variant of the original ultimatum game design first introduced by Güth et al. (1982).

<sup>15</sup>We have chosen the asymmetric ultimatum game rather than the standard ultimatum game because existing evidence on responder behavior shows that the variance in minimum acceptable offers is considerably larger in the asymmetric ultimatum game than in the standard ultimatum game. Consequently, we considered the asymmetric ultimatum game to be better suited for treatment manipulations that seek to affect responder behavior.

responders. The proposer first posts an offer of how to divide 100 chips between himself and a responder. Each responder then observes the offer and, without knowing the decision of the other responder, chooses whether or not to accept it. If both responders reject the offer, all three subjects receive zero chips. If one responder accepts the offer and one responder rejects the offer, the 100 chips are divided according to the proposed division between the proposer and the responder who accepted the offer. The responder who rejects the offer receives zero chips. If both responders accept the offer, it is randomly determined which responder actually receives the offer, and the non-selected responder receives zero chips.

In the proposer competition (PC) market game, two proposers are matched with one responder. Each proposer first posts an offer of how to divide 100 chips with the responder. The responder observes both offers and can accept one or none of the offers. If both offers are rejected, all three subjects receive zero chips. If an offer is accepted, the proposer who made the offer and the responder receive chips according to the proposed split. The proposer whose offer was not accepted receives zero chips.

### **3.2.2 Phase 2: Ultimatum Game**

In the next phase of our experiment (next 15 periods), all subjects participated in a standard version of the asymmetric ultimatum game for 15 periods. In this version, one proposer is matched with one responder. First, the proposer makes an offer to the responder. Second, the responder can accept or reject the offer. We did not elicit responders' decisions in phase 2 in the same way that we elicited them in phase 1. Before responders were informed about the actual offer, but after the offer was made, responders stated a minimum acceptable offer (MAO) amount; that is, each responder stated a number  $x$  such that the proposer's offer is accepted if and only if he offers at least  $x$  chips to the responder. This minimum amount was binding and directly enforced by the computer. As before, the proposed division of chips is implemented if and only if the proposer's offer is accepted, while both subjects get

zero chips if the proposed offer is rejected.<sup>16</sup>

A key feature of the ultimatum game that we wish to emphasize is that if the responder cares only about distributions of wealth, then his choice in the ultimatum game is completely non-strategic in the sense that it reflects only his preferences, and not his beliefs about other players' behavior. A proposer's choice in the ultimatum game, in contrast, reflects not only his preferences, but also his beliefs about the probability that the responder will accept his offer. Similarly, in the RC market game, a responder's utility from rejecting an offer may depend on whether or not he believes the other responder will accept the offer. Therefore, differences in responder behavior in the ultimatum game can be attributed to a malleability of preferences that is not captured by models of exogenously given preferences.

### 3.2.3 Procedures

At the beginning of each session, each subject was assigned to the role of proposer or responder, and this role was fixed throughout the experiment. Just before the first period, one third of the proposers and two thirds of the responders were randomly assigned to the proposer competition treatment. The remaining two thirds of the proposers and one third of the responders were assigned to the responder competition treatment. Subjects stayed in their respective treatment groups throughout all of phase 1 of the experiment. All subjects received written instructions for their respective treatment, and were asked to answer several understanding checks before proceeding with the experiment. After all subjects completed the instructions and the understanding checks, they were asked to proceed to the first phase of the experiment. Proposers and responders were randomly rematched within their treatment group after every period. The subjects were told that there would be a second phase to the experiment, but were told nothing else about it other than

---

<sup>16</sup>Our use of the strategy method in phase 2 but not in phase 1 makes the responders' choice sets very different between the two phases. In phase 1, responders are given choices  $A_1 = \{accept, reject\}$ , while in phase 2, they are given choices  $A_2 = \{0, 1, \dots, 100\}$ . We did not use this strategy in phase 1, because we didn't want to exogenously impose rules about which offer must be chosen under proposer competition. Also, note that eliciting MAOs is technically not fully equivalent to the strategy method, since a responder's full strategy might be to accept an offer of  $x$  but reject an offer  $y > x$ . But as long as responders' acceptance preferences are monotonic, there is no loss of information in eliciting MAOs.



that their choices in phase 1 would have no effect on their potential payoffs in phase 2.

Once the first phase of the experiment was finished, subjects received on-screen instructions for the ultimatum game without competition, and were again asked to work through several understanding checks. They were then divided into three different matching groups. Each matching group contained one third of the proposers and one third of the responders within a session. The first matching group consisted of proposers and responders who had previously been in the proposer competition treatment (PC Matching Group). The second matching group consisted of proposers and responders who had previously been in the responder competition treatment (RC Matching Group). Finally, the third matching group consisted of the remaining third of proposers who had previously been in the proposer competition treatment and the remaining third of responders who had previously been in the responder competition treatment (Mixed Matching Group).

As a naming convention, we will refer to responders and proposers who have previously participated in the proposer competition market as “PC Responders” and “PC Proposers”, and to those who have participated in the responder competition market as “RC Responders” and “RC Proposers”. The composition of the matching groups is summarized in table 3.1. Subjects stayed within their respective matching groups throughout all 15 periods, though the pairs were randomly reshuffled every period within each matching group.

**Table 3.1:** *Overview of Matching Groups*

	Proposer Origin	Responder Origin
PC Matching Group	PC Proposers	PC Responders
RC Matching Group	RC Proposers	RC Responders
Mixed Matching Group	PC Proposers	RC Responders

A key feature of our matching groups is that they allow us to cleanly investigate the effect of responder experience on bargaining behavior, holding proposer experience constant (by comparing the PC Matching Group with the Mixed Matching Group). Similarly, it allows us to investigate the effect of proposer experience on bargaining behavior, holding responder experience constant (by comparing the RC Matching Group with the Mixed

Matching Group).

In addition to each subject's actions, we also elicited beliefs before periods 1, 6 and 11 in both phases. Responders were asked to estimate the average offer made by proposers in the next five periods. Proposers were asked to estimate the median acceptance threshold of responders.<sup>17</sup> To avoid wealth effects potentially confounding or interfering with our treatment manipulation, either phase 1 or phase 2 was selected for payment at the end of the experiment. Within the chosen phase, 4 periods were selected at random.<sup>18</sup> The points earned in the selected periods were then converted into Swiss Francs, with the exchange rate of points to Swiss Francs set at 10:1.

In total, we ran 5 sessions totaling to 150 subjects. Because difference in past experience are a crucial variable in our design, we only invited subjects who have not previously participated in ultimatum game experiments. Sessions consisted of either 24, 30 or 36 subjects and were conducted in October and November 2012. Experiments were computerized using the software z-tree (Fischbacher, 2007) and conducted at the experimental laboratory of the University of Zurich. Our subject pool consisted primarily of students at Zurich University and the Federal Institute of Technology in Zurich.<sup>19</sup> On average, an experimental session lasted 75 minutes with an average payment of CHF 43.5 (\$47.50), including a show-up fee of CHF 15.

### 3.3 Theory and Hypothesis Development

In this section we present a simple and generally-applicable model of path-dependent fairness preferences to motivate our hypotheses.

---

<sup>17</sup>We did not incentivize the belief elicitation, and we only asked for beliefs every 5 periods for time reasons. In the proposer competition treatment, proposers were asked to additionally assume that the offer is the larger of the two offers made.

<sup>18</sup>We selected 4 periods rather than 1 to reduce the variance in subject payments in case phase 1 of the experiment was selected for payment (which otherwise would have been very large). This was necessary to comply with the payment rules of the laboratory.

<sup>19</sup>Subjects were drawn from a database of volunteers using ORSEE (Greiner, 2004).

### 3.3.1 Set-up

We begin by investigating the behavior of fairness-motivated players over the course of two phases. In phase 1, proposers and responders are matched to play either the PC market game or the RC market game. In phase 2, the players are randomly rematched to play a non-competitive ultimatum game, such that no player is matched with a player he interacted with in phase 1. In section 3.3.5 we extend the model to a setting in which subjects play for multiple periods in phase 1, and then play again for multiple periods in phase 2.

The fairness view of a player  $i$  depends on his reference point  $r_i$ , which dictates what share of the ‘total pie’ the player feels entitled to. In particular, if  $\pi_1, \pi_2, \dots, \pi_n$  are the resulting material payoffs of the players, then the utility of player  $i$  in any  $n$ -person game is given by

$$u_i(\pi_i, \pi_{-i} | r_i) = \pi_i - \lambda \max \left[ r_i(n) \left( \sum_{j=1}^n \pi_j \right) - \pi_i, 0 \right], \quad (3.1)$$

where  $\pi_{-i}$  is the vector of payoffs of all players  $j \neq i$ , and where  $r_i(n)$  denotes the share of the surplus that player  $i$  feels entitled to in an  $n$ -person game. Notice that the total pie in equation (3.1) is determined as the sum of ex-post payoffs.

We assume that, in phase 1, players begin with identical reference points  $r_i^1(n) = 1/n$  for all  $i$ . In phase 2, players update their reference points based on their phase 1 experience. Let

$$\mu_i^1 = \frac{\pi_i^1}{\sum_{j=1}^n \pi_j^1} \quad (3.2)$$

be player  $i$ 's share of the pie in phase 1, and set  $\mu_i^1 = 0$  if  $\sum_{j=1}^n \pi_j^1 = 0$ .<sup>20</sup> Then player  $i$ 's phase 2 reference point  $r_i^2$  in a  $m$ -person game is given by

$$r_i^2(m) = (1 - \gamma)(1/m) + \gamma\mu_i^1. \quad (3.3)$$

When  $\gamma = 0$ , the players' fairness preferences are not affected by their experience, and the model reduces to a simple model of *static* distributional preferences, similar to the

---

<sup>20</sup>Without having any qualitative implications for our results, one can relax this assumption to setting  $\mu_i^1 = x$  if  $\sum_{j=1}^n \pi_j^1 = 0$ , with  $x \in [0, 1/n]$ .

models introduced by Fehr and Schmidt (1999) and Bolton and Ockenfels (2000). But when  $\gamma > 0$ , players' fairness preferences are affected by their past experience, leading to a mode of *dynamic* distributional preferences. In particular, in the extreme that  $\gamma = 1$ , phase 2 preferences are completely determined by past experience.

As in our experimental design, we assume linear payoff functions. Throughout our theoretical results, we will assume that a proposer whose offer  $a$  gets accepted receives a payoff of  $k(Y - a)$  for some constant  $k > 0$ , and that the responder accepting that offer gets a payoff of  $a$ . Offers will be restricted to be in the set  $A = [0, Y]$ .

### 3.3.2 Equilibrium in the PC Market in Phase 1

We begin our analysis with the phase 1 PC market.

Our first result is that when there is proposer competition, the proposers are forced to make competitive offers at which they make zero profits. This result resembles the results of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000), who were the first to provide a theoretical account of how competitive pressures can mask agents' preferences for equal distributions.

**Proposition 17.** *In any (possibly mixed-strategy) SPE of the PC market game, at least one of the proposers offers  $a = Y$ , and the responder accepts one of the offers with probability 1. When restricting attention to pure-strategy SPEs, there is a unique equilibrium in which both proposers offer  $a = Y$ .*

The proof of Proposition 17, as well as all subsequent proofs, is contained in Appendix C.3. The intuition behind Proposition 17 is the usual Bertrand competition logic. Suppose a proposer  $i$  offers some  $a_i < Y$  and this offer is accepted with probability  $p > 0$ . Then the other proposer  $j$  certainly can't do better by offering  $a_j < a_i$ . But offering  $a_j = a_i$  is not optimal either. Proposer  $i$  can do strictly better by offering  $a_i + \epsilon$  for an arbitrarily small  $\epsilon$ , and thus increasing the probability that his offer is accepted by  $p > 0$ , while decreasing his payoff conditional on acceptance by an arbitrarily small amount.

Notice that in the PC market, the equilibrium offers are not affected by the reference transaction, and thus would not be affected by past experience. This does not mean, however, that the *responder's* minimum acceptable offer would not be affected by past experience in the PC market. The smallest offer a responder would be willing to accept in *any* subgame of the PC market game (including subgames that are off of the equilibrium path) is still determined by his reference point.

### 3.3.3 Equilibrium in the RC Market in Phase 1

Next, we characterize equilibria in the RC market game. The proposition that follows shows that in the RC market, an offer of  $a = 0$  can always be supported as an SPE outcome, though there is a multiplicity of equilibria. The multiplicity of equilibria is due to the fact that our model gives rise to a coordination game between the responders: for a wide range of parameters, it is optimal for responder  $i$  to accept an offer if and only if he thinks that the other responder will accept the offer. The source of this strategic complementarity between the responders' acceptance decisions is the behindness aversion assumed in the fairness preferences: a responder derives disutility whenever an unfair proposer gets to transact with another responder.

**Proposition 18.** *An offer  $a$  in the RC market game can be supported as an SPE if and only if*

$$a \in \left[ 0, \frac{\lambda k Y}{3 + (2+k)\lambda} \right].$$

Notice that in the RC market, a zero offers equilibrium is possible, and the highest possible offers that can be sustained in equilibrium are still significantly lower than the equilibrium offers from the PC market. In particular, even as  $\lambda \rightarrow \infty$ , the highest sustainable offers are still no larger than  $\frac{kY}{2+k}$ , meaning that proposers still get less than 1/3 of the total surplus.<sup>21</sup>

---

<sup>21</sup>In particular,  $\frac{\lambda k Y}{3 + (2+k)\lambda}$  is increasing in  $\lambda$  and  $\lim_{\lambda \rightarrow \infty} \frac{\lambda k Y}{3 + (2+k)\lambda} = \frac{kY}{2+k}$ . Now if the proposer offers  $a = \frac{kY}{2+k}$  and this offer is accepted, then his total payoff is  $2\frac{kY}{2+k}$ , from which it follows that the responder gets 1/3 fo the total sum of payoffs.

### 3.3.4 Phase 2 Behavior

To simplify exposition for the main body of the paper, we assume that in the RC market game, players coordinate on the equilibrium in which proposers offer 0.<sup>22</sup>

We begin by examining how responders' phase 2 acceptance thresholds are shaped by their phase 1 market experience. A key comparative static parameter will be  $\gamma$ —the weight that players' phase 1 experience has in shaping their phase 2 fairness preferences.

**Proposition 19.** *Let  $M_{PC}(\gamma)$  ( $M_{RC}(\gamma)$ ) be the minimal acceptable offer of a PC (RC) responder as a function of  $\gamma$ . Then*

1.  $M_{PC}(0) = M_{RC}(0)$
2.  $M_{PC}(\gamma)$  is strictly increasing in  $\gamma$
3.  $M_{RC}(\gamma)$  is strictly decreasing in  $\gamma$

The key prediction of Proposition 19 is that the higher the impact of past experience on responders' fairness reference points, the greater will be the difference between the MAOs of PC and RC responders. When no weight is given to past experience, both PC and RC responders will have identical distributional preferences, as in standard models of fairness. But as the weight  $\gamma$  increases, the MAOs of PC responders rise, while the MAOs of RC responders fall.<sup>23</sup>

We now compare the behavior of RC and PC proposers. In our next result, we examine how proposers' strategies are impacted by their Phase 1 experience and by their beliefs about the responders' behavior.

**Proposition 20.** *A PC proposer always offers  $M_{PC}(\gamma)$  to a PC responder, and offers  $M_{RC}(\gamma)$  to a RC responder. A RC proposer always offers  $M_{RC}(\gamma)$  to a RC responder. Finally, there exists a  $\gamma^\dagger \in (0, 1]$  such that*

---

<sup>22</sup>Appendix C.3 includes a more detailed analysis in which we investigate phase 2 behavior for each possible equilibrium of the RC market game. It is shown that the qualitative conclusions remain identical.

<sup>23</sup>Not that responders' behavior here is characterized by rational trading off between fairness monetary income. Zwick and Chen (1999) provide a careful experimental analysis of how subjects make tradeoffs between fairness and income.

1. If  $\gamma \leq \gamma^\dagger$  then a RC proposer always offers  $M_{PC}(\gamma)$  to a PC responder
2. If  $\gamma \in (\gamma^\dagger, 1)$  then a RC proposer offers  $a < M_{PC}(\gamma)$  to a PC responder

Proposition 20 says that PC proposers, who are used to receiving a very small share of the surplus, are predicted to act in a profit-maximizing manner and will thus offer each type of responder the smallest amount that responder is willing to accept. RC proposers, on the other hand, are used to receiving a larger share of the pie, and may not be happy with a division of surplus in which they don't get most of the pie. When an RC proposer is matched with a PC responder, there might, therefore, not be *any* division of the pie that both would find acceptable. In such a situation, an RC proposer would make an offer that would simply end in a rejection by the PC responder, as reflected in the last statement of Proposition 20. However, an RC proposer and an RC responder can always come to an agreement, since an RC responder is used to receiving very little of the surplus. In this situation, an RC proposer will make an offer equal to the smallest amount the RC responder is willing to accept.

Proposition 20 thus shows that RC proposers and PC proposers will always make identical offers to the RC responder. That is, when matched with an RC responder, proposers' Phase 1 experience has absolutely no impact on their strategy. When matched with a PC responder, RC proposers may make smaller offers than PC proposers for a high enough parameter  $\gamma$ .

Note, however, that our experimental design investigates only three of the four interactions that are considered in Proposition 20 (see Table 3.1 for a summary of the matching groups). In particular, we never match PC responders with RC proposers.<sup>24</sup> Thus Proposition 20 predicts that proposers' Phase 1 experience should not affect their strategies in our experimental design, as both types of proposers will choose strategies based solely on responders' phase 1 experience, simply offering the smallest amount their partner will be willing to accept.

---

<sup>24</sup>Since PC responders and RC proposers are always scarce in each session, such a matching group would have been prohibitively costly. Hence, we decided to focus on the other three matching groups.

### 3.3.5 Convergence

The stylized two-phase environment we have considered so far illustrates our model's predictions for how exposure to our two different markets affects subsequent behavior in the non-competitive ultimatum game. We now consider a more dynamic model in which players have the opportunity to repeatedly participate in the phase 1 market games and the phase 2 ultimatum game. The key question we ask is how behavior, and the effects of different market experiences, will change as players continue to repeatedly play the ultimatum game.

We consider play in periods  $t = -T, \dots, 0, 1, 2, \dots \infty$ . In periods  $t = -T, \dots, 0$ , players participate in some  $n$  player game (possibly one of the market games), while in periods  $t = 1, 2, \dots$  players participate in a non-competitive ultimatum game. As before, we let  $\mu_i^t$  denote the share of the pie that player  $i$  received in period  $t$ , and set  $\mu_i^t = 0$  if all  $n$  players received zero payoffs in the respective period.

In period  $t = -T$ , player  $i$ 's reference point in an  $n$  person game is given by  $\mu_i^{-T} = 1/n$ . In periods  $t > -T$ , the reference point of player  $i$  in an  $n$ -player game is given by

$$r_i^t = (1 - \gamma)(1/n) + \gamma \sum_{\tau=-T}^{t-1} \frac{\mu_i^\tau}{t + T}. \quad (3.4)$$

Notice that equation (3.4) is a generalization of equation (3.3) to a setting with more than two periods. In the more general definition (3.4), the reference point is a convex combination of the 'neutral reference point'  $1/n$  and the average of past experience. While we feel that the unweighted average of past experience is a natural input into the reference point, it is by no means the only natural specification, nor is it crucial for our results. In Appendix C.3, we show that all results remain unchanged for any weighted average of past experiences, as long as a more recent experience gets at least as much weight as an older experience.

We consider the evolution of play between a proposer and a responder in periods  $t > 0$ . We let  $r_p^t$  and  $r_R^t$  denote the proposer's and responder's period  $t > 0$  reference points. We assume that each period, proposers and responders have perfect information about each others' reference points, and play an SPE of the non-competitive ultimatum game. We let



$M^t$  denote the minimal acceptable offer of a responder  $i$  in period  $t > 0$ , and let  $a^t$  denote the proposer's period  $t > 0$  offer.

Throughout this analysis, we will be concerned with steady state preferences and strategies:

**Definition 1.** A steady state is a pair of strategies  $(a^*, M^*)$  and reference points  $(r_P^*, r_R^*)$  such that

1.  $(a^*, M^*)$  is an SPE of the ultimatum game in which players have the fairness reference points  $(r_P^*, r_R^*)$
2.  $r_P^* = (1 - \gamma)(1/2) + \gamma \frac{\pi_P^*}{\pi_P^* + \pi_R^*}$  and  $r_R^* = (1 - \gamma)(1/2) + \gamma \frac{\pi_R^*}{\pi_P^* + \pi_R^*}$ , where  $\pi_P^*$  and  $\pi_R^*$  are the proposer's and responder's steady state SPE payoffs

Our main result in this section is that there is a unique steady state to which play always converges:

**Proposition 21.** Assume that  $\gamma < 1$ . Then there is a unique steady state  $\langle (a^*, M^*), (r_P^*, r_R^*) \rangle$ . In the steady state,  $a^* > 0$ ,  $a^* < k(W - a^*)$ , and  $a^* = M^*$ . Moreover, this steady state is globally stable. That is, for any set of initial experiences  $\{\mu_i^t\}_{t=-T}^0$ , preferences and strategies converge to the steady state:

$$\begin{aligned} \lim_{t \rightarrow \infty} r_P^t &= r_P^* & \text{and} & & \lim_{t \rightarrow \infty} r_R^t &= r_R^* \\ \lim_{t \rightarrow \infty} a^t &= a^* & \text{and} & & \lim_{t \rightarrow \infty} M^t &= M^* \end{aligned}$$

Proposition 21 shows that if players have enough experience in the ultimatum game environment, then their fairness preferences in that environment can be characterized as a fixed point of an adjustment dynamic. In fact, Proposition 21 shows that our model uniquely pins down what the steady-state fairness preferences can be—the steady state is unique. The only assumption needed to guarantee uniqueness is that  $\gamma < 1$ : that is, that players' fairness preferences are not completely (though perhaps arbitrarily close to) determined by past experience.

A final prediction of the model is that when players have extreme past experiences as in our market conditions, convergence to the steady state will be monotonic. That is,

PC responders should monotonically decrease their MAOs, while RC responders should monotonically increase their MAOs:

**Proposition 22.** *Assume that  $\gamma < 1$  and that  $\frac{\sum_{i=-T}^0 \mu_R^i}{T+1} + \frac{\sum_{i=-T}^0 \mu_P^i}{T+1} \leq 1$ .*

1. *If  $\frac{\sum_{i=-T}^0 \mu_R^i}{T+1} < r_R^*$ , then for all  $t > 0$ ,  $M^t < r_R^*$  but is strictly increasing in  $t$ .*
2. *If  $\frac{\sum_{i=-T}^0 \mu_R^i}{T+1} > r_R^*$ , then for all  $t > 0$ ,  $M^t > r_R^*$  but is strictly decreasing in  $t$ .*

Proposition 22 simply says that even though responders' MAOs should not reach steady state levels in a finite number of periods, the effect of past market experience should still diminish over time.

### 3.3.6 Discussion of Assumptions

To keep the analysis as simple and clear as possible, we make several simplifying assumptions in the model that are almost surely at odds with behavior in our experiment. In particular, we make the extreme assumption that players derive disutility when their share of the pie falls short of their reference point, but they do not derive disutility from receiving a disproportionately large share of the pie. Incorporating such a motive would weaken the statement of Proposition 20. However, as long as proposers derive greater disutility from falling short of their reference point than from exceeding it, the qualitative prediction of Proposition 20 will still hold: that is, proposers' strategies will depend more on responders' Phase 1 experience than on their own Phase 1 experience.

### 3.3.7 Testable Hypotheses

Our theoretical results lead to a number of testable hypotheses. We begin by enumerating the hypotheses for Phase 1 of the experiment.

**H1** In the PC market, proposers will offer most of the pie. In the RC market, proposers will offer significantly less than proposers in the PC market.

The first part of hypothesis H1 is a direct consequence of Proposition 17: because of Bertrand-style competition, proposers cannot maintain positive material surplus. The second part of hypothesis H1 is a consequence of Proposition 18. While proposition 18 does not pin down a unique equilibrium, it does say that even the largest possible equilibrium offer is far below the predicted equilibrium in the PC market, and it says that even offers of zero can be supported as an equilibrium in the RC market.

We next turn to hypotheses for Phase 2 of the experiment.

**H2.1** Responders from the PC market will have higher MAOs than responders from the RC market

**H2.2** Proposers who have learned from Phase 2 experience what responders in their matching group are willing to accept will offer more to PC responders than to RC responders.

**H2.3** Proposers who have learned from Phase 2 experience what responders in their matching group are willing to accept will not be affected by their Phase 1 market experience.

**H2.4** Differences in Responders' strategies due to Phase 1 experience will diminish over time, but will not be completely eliminated.

Hypothesis H2.1 is formally derived in Proposition 19, and is a basic consequence of assuming that preferences are path-dependent.

Hypothesis H2.2 is derived in Proposition 20, and is a consequence of Proposers' profit-maximization motives: proposers should offer less to responders who are willing to accept less. We expect H2.2 to be in full force after a few periods of play, once Proposers have a chance to learn what offers responders are willing to accept.

Hypothesis H2.3, also motivated by Proposition 20, complements H2.2 and says that proposers' own experience in Phase 1 should play no role in their strategies once they have an opportunity to learn about responders' preferences. Initially, it is possible that proposers' Phase 1 experience may affect their beliefs about responder behavior, and thus affect their strategies. But when proposers learn about responders' strategies, Proposition 4's prediction that proposers' Phase 1 experience will not affect their strategies should be in full force.

Hypothesis H2.4 is directly motivated by Proposition 22, which states that with experience, proposers' reference points should monotonically converge toward the steady-state reference point.

## 3.4 Results

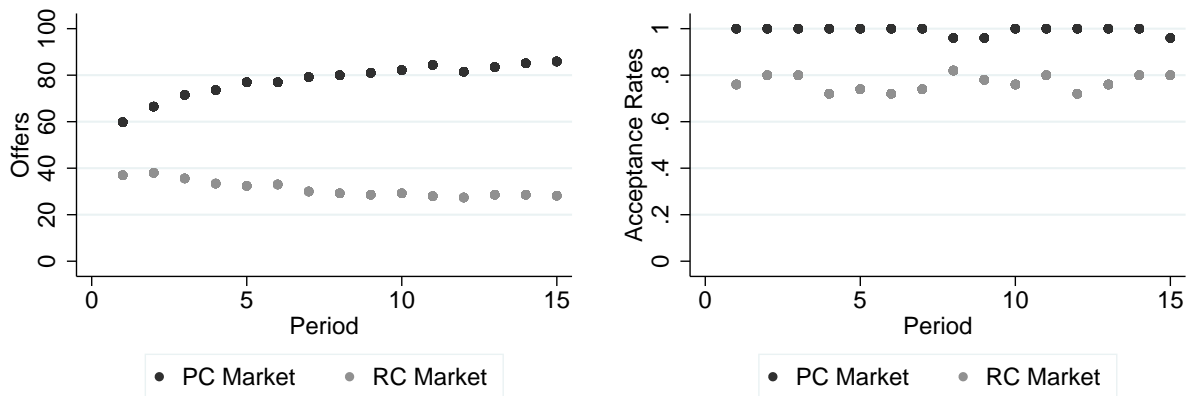
In this section, we present our experimental evidence for path-dependent fairness preferences. We begin in subsection 3.4.1 by analyzing differences in behavior between the RC and PC markets: as expected, we find that offers in the PC market are significantly higher than offers in the RC market. In subsection 3.4.2 we turn to the analysis of phase 2 of the experiment, and investigate how differences in phase 1 experience affect responder behavior in phase 2. In subsection 3.4.3 we investigate how differences in phase 1 experience affect proposers' offers. Finally, in subsection 3.4.4 we investigate to what extent responders' fairness preferences converge over time.

### 3.4.1 Phase 1: The Effect of Competition on Offers and Acceptance Decisions

We find strong evidence that competition affects offers in the first phase of our experiment. Averaged over all 15 periods, proposers offer 78 chips to responders in the PC market, whereas they offer only 31 chips to responders in the RC market. The development of offers over the course of the 15 periods in both treatments is shown in the left panel of figure 3.1. The difference between offers in the two treatments is roughly 23 chips in period 1, and increases over time until it reaches an average of 50 chips from period 7 onwards. To assess the statistical significance of this difference, we use a clustered version of the rank-sum test proposed by Datta and Satten (2005), which controls for potential dependencies between observations. Clustering on the treatment group in phase 1, we find that the difference in individual average offers is highly significant ( $p < 0.01$ , clustered rank sum test)<sup>25</sup>

---

<sup>25</sup>While we present clustered rank-sum tests in the text, we also present p-values using standard rank sum tests that do not account for the potential dependencies between observations within matching groups in table C.1 in Appendix A as a benchmark. It can be seen that the p-values of the standard test are in general slightly smaller, but very comparable to the p-values from the clustered test presented in the text.



**Figure 3.1:** Average Offers and Acceptance Rates Over Time, Respectively.

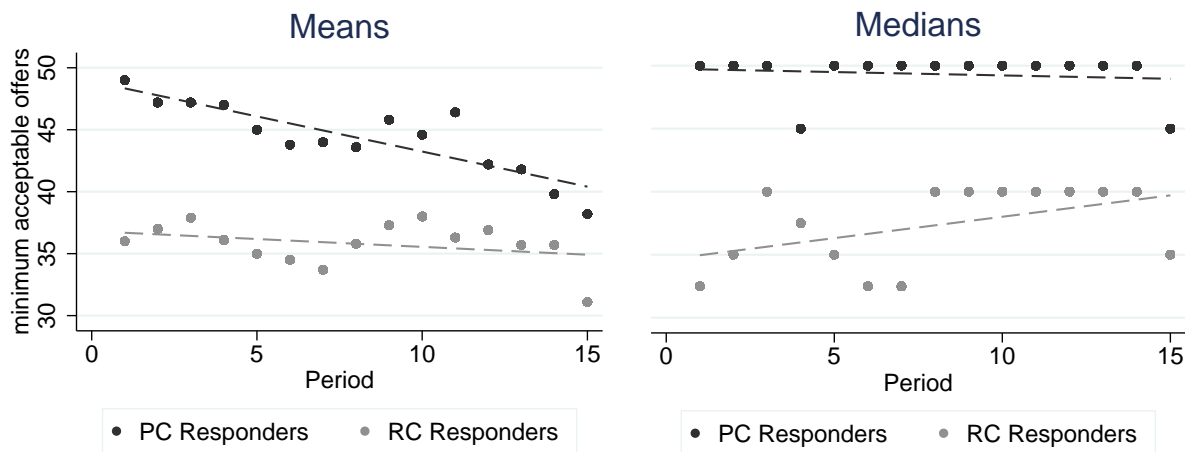
Despite these large differences in offers, the right panel of figure 3.1 shows that the probability that an offer is accepted does not differ much by treatment. In the PC market, responders accept one of the two offers in 99.2 percent of the time. In the RC market, responders accept the offers 76.8 percent of the time, and the probability that *at least one* of the responders accepts an offer is 92.5 percent. Thus in both markets, a successful transaction occurs over 90 percent of the time. Our stark experimental results on the affects of competitive forces are consistent with Roth et al. (1991) and Fischbacher et al. (2009) and confirm our hypothesis 1. We summarize them in the following result:

**Result 1.** *Competition among responders leads to low offers, whereas competition among proposers leads to high offers in the market version of the ultimatum game. Despite facing relatively low offers, responders in the responder competition market accept offers with a high frequency.*

### 3.4.2 Phase 2: The Effect of Experience on Responder Behavior

Hypothesis H2.1 states that PC responders have higher acceptance thresholds than RC responders. The left panel of figure 3.2 provides our first piece of evidence in favor of the hypothesis. In every period of the bargaining game, average minimal acceptable offers are larger for PC Responders. The difference is particularly pronounced in early periods. In period 1, the difference in the average acceptance threshold between the two

treatment groups is 13 chips, which translates to PC Responders stating minimum acceptance thresholds that are 36 percent higher than the acceptance thresholds of RC Responders.



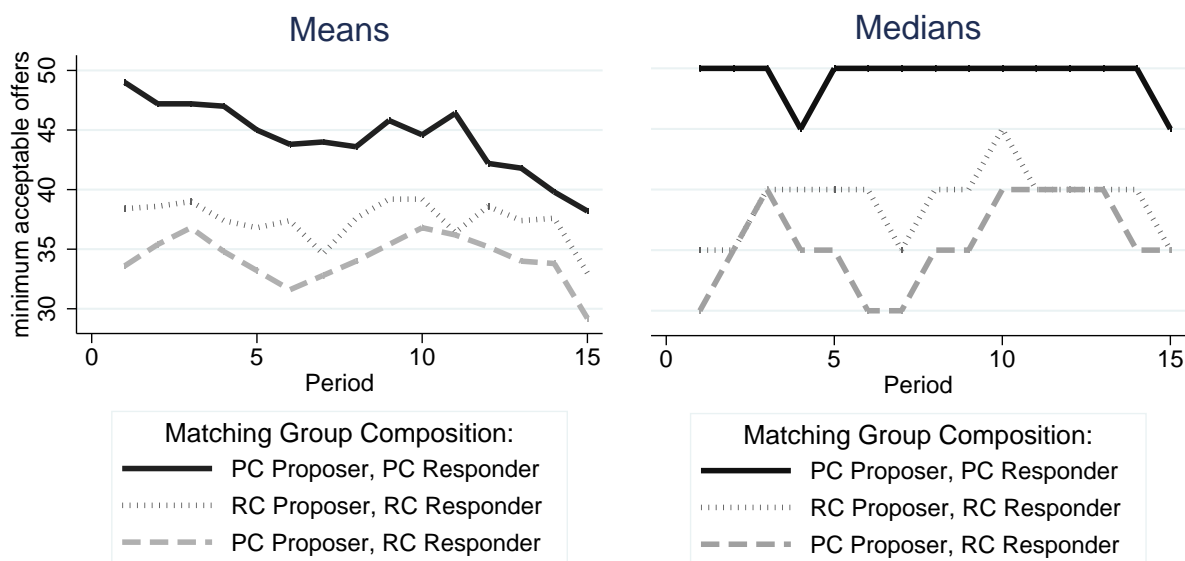
**Figure 3.2:** *Minimal Acceptance Thresholds of Responders.*

To assess the statistical significance of this difference, we again use the clustered version of the rank-sum test in order to control for potential dependencies between observations stemming from phase 1 of the experiment. Clustering on the treatment group in phase 1, the difference in minimum acceptance thresholds in the first period is significant ( $p = 0.03$ , clustered rank-sum test). On average over all 15 periods, the difference is 8.6 chips ( $p = 0.06$ , clustered rank-sum test using individual average acceptance thresholds as the unit of observation), which translates to PC responders stating minimum acceptance thresholds that are 24 percent higher.

The right panel of figure 3.2 shows that a similar picture emerges when comparing the median minimum acceptance thresholds across the two treatment groups. The median minimum acceptance threshold of the PC responders is consistently larger than the median minimum acceptance threshold of the RC responders, and the treatment effect is very consistent over time. A ranksum test using median acceptance thresholds of each treatment group in each session as the unit of observation ( $N = 10$ ) reveals that the difference in medians is also significant ( $p < 0.02$ ).

The data shown in figure 3.2, however, does not account for differences in past experience

of the proposer. To control for potential effects of proposer experience on responder behavior, we have to exploit the exogenous variation in matching group composition in phase 2 of the experiment. As we explained in section 3.2, our different matching groups in phase 2 of the experiment allow us to perform such an analysis. Figure 3.3 shows the development of minimum acceptance thresholds over time in phase 2 of the experiment for the three different matching groups. While the solid line exactly corresponds to the PC responder plot in figure 3.2, the two dotted lines disaggregate the RC responders based on their matching groups. It can again be seen that minimum acceptance thresholds in the PC matching group remain permanently above the minimum acceptance thresholds of the other two groups. Moreover, the two matching groups with RC responders are not significantly different from each other ( $p = 0.44$ , clustered rank-sum test using individual average acceptance thresholds as the unit of observation).



**Figure 3.3:** Mean and Median Minimal Acceptance Thresholds of Responders, by Matching Group.

To statistically assess potential differences in responder behavior conditional on responder and proposer experience, we turn to regression analysis. Table 3.2 shows coefficients of a regression of minimum acceptance thresholds on a dummy variable indicating whether a

responder participated in the PC market in phase 1 (PC Responder) and a dummy variable indicating whether the matched proposer participated in the PC market in phase 1 (PC Proposer). Hence, PC Responder captures the effect of responder experience, whereas PC Proposer captures the effect of proposer experience.

**Table 3.2:** *Minimum acceptable offers by experience*

	First Period	All
	(1)	(2)
PC Responder	15.400** (5.665)	10.187* (4.921)
PC Proposer	-4.800 (6.777)	-3.227 (7.490)
Constant	38.400*** (4.165)	37.413*** (4.675)
Adj. $R^2$	0.059	0.041
Observations	75	1125

OLS Regressions; Clustering by treatment groups. Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

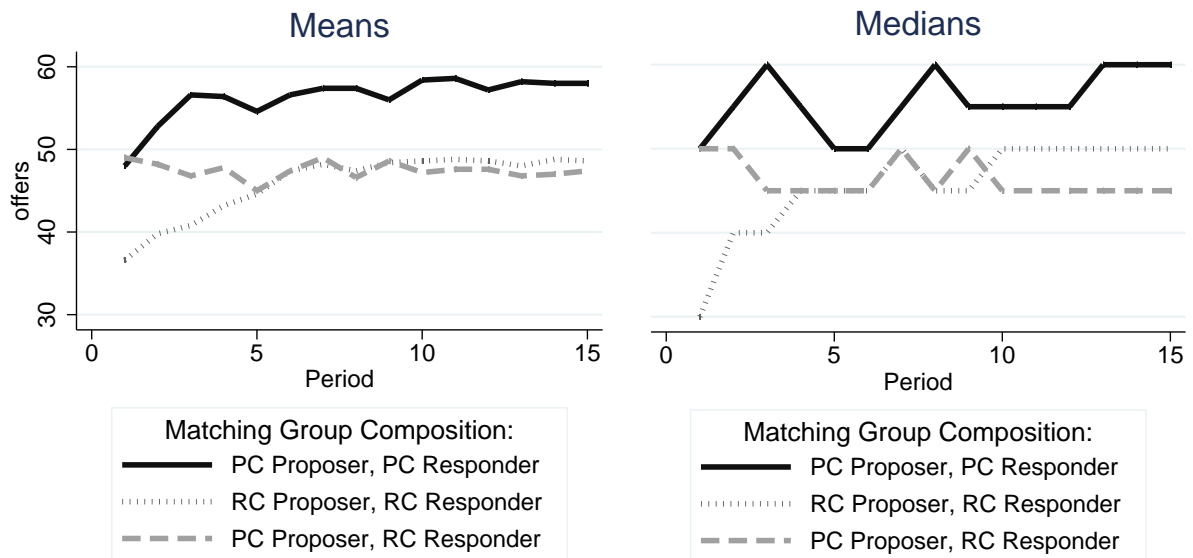
Table 3.2 clearly shows that only responder experience has a significant effect. This effect is particularly pronounced in the first period, and weakens when all periods are considered in the regression. We will return to the issue of convergence in subsection 3.4.4. Our evidence on responder behavior is summarized in the next result:

**Result 2.** *Responders who have been exposed to proposer competition in phase 1 have a higher minimal acceptance threshold than responders who have been exposed to responder competition in phase 1. The experience of matched proposers has no substantial impact on responder behavior.*

### 3.4.3 The Effect of Experience on Proposer Behavior

Our theoretical framework also makes predictions about proposer behavior. Hypothesis H2.2 states that proposers should tailor their phase 2 offers to the responders' past phase 1 experience. Hypothesis H2.3 states that once proposers have an opportunity to learn about responders' behavior, proposers should not be affected by their own phase 1 experience at all. We now investigate these hypotheses.





**Figure 3.4:** Mean and Median Offers of Proposers, by Matching Group.

The left panel of figure 3.4 shows average offers in the different matching groups over time, and the right hand panel shows the respective median offers by matching group over time. It can immediately be seen that, in the first period, proposer origin affects proposer offers. While the mean and median offer is roughly equivalent in the two matching groups with PC proposers, RC proposers make lower offers in period 1. However, mean and median offers of PC proposers seem to diverge, whereas mean and median offers of the RC proposers appear to converge to the level of the PC proposers who are matched with RC responders. To assess the relevance of proposer and responder experience on proposers' offers, we again turn to regression analysis.

Table 3.3 shows a regression of offers on a dummy variable indicating whether a responder participated in the PC market in phase 1 (PC Responder) and a dummy variable indicating whether the matched proposer participated in the PC market in phase 1 (PC Proposer). Column (1) shows that proposers' phase 1 experience is an important determinant of their first period offers. When aggregating over all 15 periods, however, the effect is completely reversed. The coefficient on PC Proposer becomes insignificant, whereas the

**Table 3.3: Proposer Offers**

	First Period	All
	(1)	(2)
PC Responder	-1.000 (3.835)	8.813* (3.927)
PC Proposer	12.400*** (2.461)	1.613 (2.541)
Constant	36.600*** (2.836)	45.853*** (3.049)
Adj. $R^2$	0.068	0.106
Observations	75	1125

Notes: OLS Regressions; Clustering by treatment groups. Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

coefficient on PC Responder becomes larger and significant ( $p = 0.051$ ). Thus, once proposers have learned the preferences of the responders with whom they are matched, their phase 1 experience becomes completely irrelevant. We summarize these findings in the following result:

**Result 3** (Offers are Primarily Driven by Responder Experience). *Overall, proposers' experience in phase 1 of the experiment has very little impact on their offers in phase 2 of the experiment. Responders' phase 1 experience, however, significantly impacts proposers' offers in phase 2 of the experiment.*

Data on proposers' beliefs further supports the conclusion that proposers' offers in phase 2 are strongly affected by their beliefs about responders' minimum acceptance thresholds in phase 2.<sup>26</sup> In Periods 1, 6 and 11, proposers were asked to state their belief about responders' median acceptance threshold. Table 3.4 shows regressions of this believed median acceptance threshold on proposer and responder experience dummies.

PC Proposers initially believe in higher acceptance thresholds, but the difference is not statistically significant. However, in period 6 and 11 the effect of proposer experience on believed acceptance thresholds vanishes. Responder experience, however, increasingly

<sup>26</sup>This is consistent with earlier work documenting that bargainers' beliefs about their partners' strategies will effect their own strategies. See, e.g., Roth and Schoumaker (1983).

**Table 3.4:** *Proposers' beliefs about acceptance*

	Period 1	Period 6	Period 11
	(1)	(2)	(3)
PC Proposer	6.600 (9.995)	0.600 (4.264)	-4.400 (5.237)
PC Responder	6.800 (10.449)	12.000* (4.341)	15.800*** (2.736)
Constant	35.000*** (4.657)	44.000*** (3.463)	46.800*** (4.558)
Adj. $R^2$	0.050	0.149	0.199
Observations	75	75	75

Notes: OLS Regressions; Clustering by treatment groups. Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

affects the believed median acceptance threshold. While it initially has no effect in period 1, being matched with a PC responder leads to significantly higher beliefs in median acceptance thresholds in periods 6 and 11. This evidence is consistent with proposers offering higher amounts when being matched with PC responders.

The fact that proposers' offers are strongly shaped by their beliefs about responder behavior is also consistent with our theoretical predication that proposer behavior is driven by the maximization of monetary payoffs. To further investigate this prediction, we compute the optimal offers for each period (and by matching group) in appendix C.2, and find that proposers' actual offers are very close to the optimal offers (both in magnitudes and statistically).

#### **3.4.4 Convergence of Fairness Preferences**

Next, we turn to the persistence of the effects of experience on proposers and responders over time. According to Hypothesis 2.4, the difference between PC and RC responders' MAOs should diminish over time. In this section, we analyze how quickly this convergence occurs, if at all. The fitted lines in figure 3.2 suggest that there is a negative trend in minimum acceptable offers for responders who have been exposed to proposer competition in the first part of the experiment. This observation is confirmed by regression analysis in

column (1) of table 3.5.

**Table 3.5:** *Time trends in minimum acceptance thresholds*

	(1)
PC Responder	12.093** (4.296)
PC Responder × period	-0.440* (0.222)
Period	-0.126 (0.182)
Constant	36.811*** (2.449)
Adj. $R^2$	0.041
Observations	1125

*Notes:* OLS regressions; Clustering by treatment groups

The difference in acceptance thresholds between the two treatment groups is decreasing by an average of .44 chips per period, which is approximately 3.6 percent of the initial difference. This attenuation of the difference appears to be mainly driven by responders from the PC treatment group slowly decreasing their acceptance thresholds over time (see the significant negative time trend in column (1)). Responders from the responder competition treatment group, however, do not change their acceptance thresholds over time. We summarize these findings in our final Result:

**Result 4** (Gradual Convergence of Fairness Preferences). *The difference in acceptance thresholds is fairly persistent, but gradually decreasing over time.*

### 3.4.5 Discussion of Results

#### Do Experiences or Expectations Shape the Reference Point?

In the theoretical development of our hypotheses, we have assumed that fairness reference points correspond to past experiences. Alternatively, following the approach by Kőszegi and Rabin (2006, 2007, 2009) in the domain of consumption and risk preferences, it is possible that reference points correspond to expectations. Applied to fairness preferences, it may be

that a responder chooses to reject a proposer's offer when that offer falls far short of what the responder expected.<sup>27</sup>

A model of *rational* expectations along the lines of Kőszegi and Rabin (2006, 2007, 2009) would not predict our results. Once players learn which game they will be participating in for the subsequent 15 periods, their rational expectations about outcomes should not depend on their phase 1 experiences.<sup>28</sup>

Another possibility is that a model of *naive*, rather than rational, expectations might be generating the differences in responders' phase 2 behavior. Although one is faced with many degrees of freedom in formulating a generally-applicable theory of naive expectations that can explain how beliefs in one game are shaped by outcomes in a previous game, one of the many possible formulations is similar to our model. A model in which players' expectations are a convex combination of the shares of the surplus that they have received in the past could generate patterns similar to the patterns observed in our data. In that sense, our model can be viewed as a tractable and reduced-form approach to formalizing the behavior of a player whose experiences shape his naive beliefs, which in turn affect preferences.<sup>29</sup>

We want to point out, however, that a model in which naive beliefs shape the fairness reference point would have trouble explaining the stark difference between responders' adjustment of their phase 2 behavior and proposers' adjustment of their phase 2 behavior (see Figures 3.3 and 3.4). As we discussed in detail, proposer behavior suggests that our

---

<sup>27</sup>In the context of third-party punishment, Coffman (2010) tests the idea that third parties' expectations may shape their punishment decisions, but does not find evidence for this hypothesis.

<sup>28</sup>Rational expectations should only be shaped by knowledge of the game structure, and beliefs about other players' types. And since rational players should not have their beliefs systematically biased by play in different games, these rational players should not have different beliefs about each others' types as a result of playing different games in Phase 1. Of course, it may be possible to *accommodate* our results with a model in which there are multiple rational expectations equilibria and past experience serves as a coordination device for selecting an equilibrium. However, we do not find such an explanation particularly satisfactory, since it amounts to assuming a model with enough degrees of freedom in its predictions such that our data can't falsify it. A more satisfactory account would have our empirical results as a *prediction*.

<sup>29</sup>A responder's naive beliefs might literally be dictated by a his past shares of the surplus—thus corresponding exactly to our mathematical model—or the beliefs might be a convex combination of the offers the responder observes—thus corresponding to a minor variation of our model that generate identical results.

subjects are actually very fast to learn how other subjects behave in the game they are playing. By period 5, proposers' phase 1 experience has absolutely no impact on their behavior. Given such fast learning, a theory in which differences in responders' behavior are due solely to differences in their naive expectations would seem to be mismatched with the persistent difference in MAOs that we observe in our data (see Figure 3.3).<sup>30</sup> Finally, although reinforcement learning (Roth and Erev, 1995, 1998) does not make predictions how experience in one game would modify behavior in a different game with a different strategy space, like our model it also predicts that responders may adjust their behavior in response to the offers they see (Cooper et al., 2003).<sup>31</sup>

### **Fairness Reference Points or Anchoring Heuristics?**

Experimental evidence has shown that individuals can be influenced by arbitrary anchors (Lichtenstein and Slovic, 2006; Kahneman and Tversky, 2000; Ariely et al., 2003; Simonson and Tversky, 1992), and that behavior that appears to be consistent with expressing a particular preference can in fact be the result of arbitrary anchoring.<sup>32</sup> Could it then be that our results are the consequence of a simple anchoring heuristic rather than evidence for reference-dependent fairness preferences?

Two pieces of evidence suggest that our results are not due to simple anchoring. First, we demonstrate that past experience only has persistent effects on responder behavior, whereas proposers quickly adapt to the new environment. This differential response to the treatment

---

<sup>30</sup>We should also point out that while a theory of expectations as reference points is the most similar to our model, similar considerations apply to belief-based reciprocity models (Rabin, 1993; Levine, 1998; Dufwenberg and Kirchsteiger, 2004). In these models players' motivations depend on their second-order beliefs about actions (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004) or on first-order beliefs about types (Levine, 1998). Again, a theory of reciprocity and *rational* beliefs would not be consistent with our results. But when combined with certain types of non-rational beliefs, these theories may come closer to rationalizing our results. Cooper and Dutcher (2011) propose a model along these lines to explain their findings from a meta-analysis of responder behavior in ultimatum games.

<sup>31</sup>See also List and Cherry (2000) and Slonim and Roth (1998) for work on proposer and responder adjustment in the ultimatum game.

<sup>32</sup>See, however, Fudenberg et al. (2012) and List et al. (2013) for evidence questioning the robustness of these anchoring effects.

is inconsistent with a simple anchoring heuristic, whereas our theory of reference-dependent fairness preferences precisely predicts such behavior.

Second, we collected data on individual cognitive abilities using the cognitive reflection test (CRT) (Frederick, 2005). The CRT is a cognitive test that seems particularly suited to assess the proneness to the anchoring heuristic. In the CRT, subjects have to answer questions that have an intuitive but wrong answer. Individuals who fall for fast and frugal heuristics are likely to score low on the CRT. Table C.2 in appendix A shows coefficients for regressions of the minimum acceptance threshold on phase 1 experience, CRT scores and the interaction of the treatment with the CRT score in period 1 of phase 2 of the experiment, in which anchoring should be particularly pronounced. If our subjects were following a simple anchoring heuristic, we would expect that the treatment effect is particularly pronounced for individuals scoring low on the CRT. It turns out that the statistically insignificant point estimate of the interaction between CRT score and treatment has the opposite sign.

### **3.5 Concluding Remarks**

While most existing work on social preferences has progressed under the presumption of static preferences, we show that fairness preferences are malleable and endogenous to the economic forces that determine market outcomes. We also demonstrate that such malleability can be incorporated into economic theory in a disciplined generalizable way: the parsimonious and generally applicable model we propose incorporates the path-dependent nature of preferences at the cost of just one additional degree of freedom, and predicts uniquely determined long-run outcomes in the games we consider.

We view our experimental and theoretical work as a first step towards understanding the dynamic nature of fairness preferences, paving the way for a number of other theoretical and empirical questions to be addressed by future research. One important set of questions concerns the foundations of path dependence. Although our results would not be predicted by a model in which agents form rational beliefs and derive utility from those beliefs, our mathematical framework and experimental results are not meant to distinguish between

whether experience is a direct input into preferences, or whether the input is beliefs that are shaped by the past experiences of boundedly-rational agents. We plan to investigate this question in future work.

While past experience does not appear to affect proposers' prosociality in our empirical results, a relationship between prosociality and past experience may exist in other contexts. In concurrent work, Peysakhovich and Rand (2013) demonstrate that immersing subjects in environments that support cooperation in the infinitely repeated Prisoner's Dilemma (PD) substantially impacts subjects' norms of prosociality in subsequent interactions. Peysakhovich and Rand (2013) find that subjects who have previously experienced cooperative outcomes are more likely to cooperate across a broad array of one-shot games such as the dictator and trust games, and are also more likely to punish selfishness in third-party punishment games. Interestingly, Peysakhovich and Rand (2013) find no effect on retaliation in the ultimatum game, which suggests that there may be important nuances in how past experience shapes retaliatory versus cooperative motivations. We believe this to be an important theoretical and empirical question for future research.<sup>33</sup>

Our work also raises intriguing questions about the market consequences of dynamic fairness preferences. For example, what are the implications of path-dependent fairness preferences for how a firm would optimally choose its dynamic price schedule? Our work implies that a trade-off exists between the immediate loss of customers whose fairness reference point is violated, and the long run profits generated through the increased willingness to pay of customers once the reference point has adjusted. And unlike theories of reference dependence that do not invoke fairness preferences—but in line with the insights of Kahneman et al. (1986)—our framework implies that consumers will react very differently to price increases that are exploitations of market power, as opposed to price increases necessitated by rising costs of production.

---

<sup>33</sup>Also quite intriguing is that Peysakhovich and Rand (2013) find that subjects with lower CRT scores are most affected by their PD manipulation, which leads them to suggest that subjects adopt cooperative tendencies as "heuristics" shaped by past experiences. In contrast, we find no interaction between CRT scores and the effect of past experience. This suggests that there may be fundamental differences between cooperative and retaliatory behaviors, and that past experiences may shape these behaviors through different psychological channels.



The effect of experience on fairness perceptions in labor markets may also cause wage rigidities and create excessive unemployment volatility. Finally, the path-dependence of fairness preferences may also help to shed light on the differences in beliefs and attitudes that we observe across different cultures and institutions.

# References

- Acland, Dan and Matthew Levy, "Habit Formation, Naivete, and Projection Bias in Gym Attendance," *working paper*, 2011.
- Agarwal, Sumit, John C. Driscoll, Xavie Gabaix, and David Laibson, "Learning in the Credit Card Market," *working paper*, 2013.
- Akerlof, George A, "Labor contracts as partial gift exchange," *The Quarterly Journal of Economics*, 1982, 97 (4), 543–569.
- Allcott, Hunt. "The Welfare Effects of Misperceived Product Costs: Data and Calibrations from the Automobile Market," *American Economic Journal: Economic Policy*, 2013, 5 (3) 30–66.
- Allcott, Hunt, and Michael Greenstone, "Is there an Energy Efficiency Gap?" *Journal of Economic Perspectives*, 2012, 26 (1) 3–28.
- Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky, "Energy Policy with Externalities and Internalities," *Journal of Public Economics*, forthcoming.
- Allcott, Hunt and Todd Rogers, "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," *NBER Working Paper 18492*, 2012.
- Allcott, Hunt, and Nathan Wozny, "Gasoline Prices, Fuel Economy, and the Energy Paradox," *Review of Economics and Statistics*, 2012, forthcoming.
- Altmann, Steffen and Christian Traxler, "Nudges at the Dentist," *working paper*, 2012.
- Anderson, Dennis, and John Claxton, "Barriers to Consumer Choice of Energy Efficient Products," *Journal of Consumer Research*, 1982, 9 (2) 163–170.
- Anderson, Soren, Ryan Kellogg, and James Sallee. What Do Consumers Believe about Future Gasoline Prices? NBER Working Paper No. 16974, 2011.
- Andreoni, James and Charles Sprenger, "Estimating Time Preferences from Convex Budgets," *American Economic Review*, 2012, 102 (7), 3333–3356.
- Ariely, Dan and Klaus Wertenbroch, "Procrastination, Deadlines, and Performance: Self-Control by Precommitment," *Psychological Science*, 2002, 13 (3), 219–224.

- , George Loewenstein, and Drazen Prelec, ““Coherent arbitrariness”: Stable demand curves without stable preferences,” *The Quarterly Journal of Economics*, 2003, 118 (1), 73–106.
- Armstrong, April W., Alice J. Watson, Maryanne Makredes, Jason E. Frangos, Alexandra Kimball, and Joseph C. Kvedar, “Text-Message Reminders to Improve Sunscreen Use,” *Archives of Dermatology*, 2009, 145 (11), 1230–1236.
- Attari, Shahzeen, Michael DeKay, Cliff Davidson, and Wandu Bruine de Bruin, “Public Perceptions of Energy Consumption and Savings,” *Proceedings of the National Academy of Sciences*, 2010, 107 (37) 16054–16059.
- Atkinson, Richard C. and Richard M. Shiffrin, “Storage and Retrieval Processes in Long-Term Memory,” *Psychological Science*, 1969, 76 (2), 179–193.
- Baicker, Kate, Sendhil Mullainathan, and Joshua Schwartzstein, “Choice Hazard in Health Insurance,” Working Paper, Dartmouth College, 2013.
- Banks, Jeffrey and Sridhar Moorthy, “A Model of Price Promotions with Consumer Search,” *International Journal of Industrial Organization*, 1999, 17, 371–398.
- Bartling, Björn and Klaus M. Schmidt, “Reference points in renegotiations: The role of contracts and competition,” ECON - Working Papers 089, Department of Economics - University of Zurich 2012.
- Beaudry, Paul and John DiNardo, “The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data,” *Journal of Political Economy*, 1991, pp. 665–688.
- Becker, Gary, and Kevin Murphy, “A Simple Theory of Advertising as a Good or Bad,” *Quarterly Journal of Economics*, 1993, 108 (4) 941–964.
- Becker, Gary S. and Kevin M. Murphy, “A Theory of Rational Addiction,” *The Journal of Political Economy*, 1988, 96 (4), 675–700.
- and – , “A Simple Theory of Advertising as a Good or Bad,” *Quarterly Journal of Economics*, 1993, 108 (4), 941–964.
- Bednar, Jenna, Yan Chen, Tracy Xiao Liu, and Scott Page, “Behavioral spillovers and cognitive load in multiple games: An experimental study,” *Games and Economic Behavior*, 2012, 74 (1), 12–31.
- Benjamin, Daniel J., “A Theory of Fairness in Labor Markets,” *Working paper*, 2005.
- Bernheim, Douglas B. and Antonio Rangel, “Addiction and Cue-Triggered Decision Processes,” *The American Economic Review*, 2004, 94 (5), 1558–1590.
- Bernheim, B. Douglas, and Antonio Rangel, “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics,” *Quarterly Journal of Economics*, 2009, 124 (1) 51–104.

- Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman, "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment," *Quarterly Journal of Economics*, 2010, 125 (1), 1–45.
- Bewley, Truman F., *Why Wages Don't Fall During A Recession*, Cambridge, MA: Harvard University Press, 1999.
- Bhargava, Saurabh, and Dayanand Manoli, "Why Are Benefits Left on the Table? Assessing the Role of Information, Complexity, and Stigma on Take-up with an IRS Field Experiment," Working Paper, University of Texas at Austin, 2013.
- Binmore, Ken, Peter Morgan, Avner Shaked, and John Sutton, "Do people exploit their bargaining power? An experimental study," *Games and Economic Behavior*, 1991, 3 (3), 295–322.
- Blumstein, Carl, Betsy Krieg, Lee Schipper, and Carl York, "Overcoming Social and Institutional Barriers to Energy Conservation," *Energy*, 1980, 5 355–371.
- Bollinger, Brian, Phillip Leslie, and Alan Sorensen, "Calorie Posting in Chain Restaurants." *American Economic Journal: Economic Policy*, 2011, 3 (1) 91–128.
- Bolton, Gary E. and Axel Ockenfels, "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 2000, 90 (1), 166–93.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," *working paper*, 2013.
- Bureau of Labor Statistics (BLS), "Consumer Expenditure Survey: Expenditure Tables," 2013. <http://www.bls.gov/cex/csxstnd.htm>.
- Borenstein, Severin, and Lucas Davis, "The Equity and Efficiency of Two-Part Tariffs in U.S. Natural Gas Markets," *Journal of Law and Economics*, 2012, 55 (1).
- Busse, Meghan, Christopher Knittel, and Florian Zettelmeyer, "Are Consumers Myopic? Evidence from New and Used Car Purchases." *American Economic Review*, 2013, 103 (1) 220–256.
- Butters, Gerard R., "Equilibrium Distributions of Sales and Advertising Policies," *The Review of Economic Studies*, 1977, 44 (3), 465–491.
- Cabral, Marika, and Caroline Hoxby (2012), "The Hated Property Tax: Salience, Tax Rates, and Tax Revolts," NBER Working Paper 18514, 2012.
- Cadena, Ximena and Antoinette Schoar, "Remembering to Pay? Reminders vs. Financial Incentives for Loan Repayments," *NBER working paper #17020*, 2011.
- Calzolari, Giacomo and Mattia Nardotto, "Nudging with Information: A Randomized Field Experiment," *working paper*, 2012.
- Camerer, Colin, "Progress in Behavioral Game Theory," *Journal of Economic Perspectives*, 1997, 11, 167–188.

- , *Behavioral Game Theory - Experiments in Strategic Interaction*, Princeton, New Jersey: Princeton University Press, 2003.
- Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue, and Matthew Rabin, "Regulation for Conservatives: Behavioral Economics and the Case for Asymmetric Paternalism," *University of Pennsylvania Law Review*, 2003, 151 1211–1254.
- Cameron, Colin, Jonah Gelbach, and Douglas Miller, "Robust Inference with Multi-way Clustering," *Journal of Business and Economic Statistics*, 2011, 29 (2), 238–249.
- Caplin, Andrew and Mark Dean, "The Behavioral Implications of Rational Inattention with Shannon Entropy," *working paper*, 2013.
- and – , "Rational Inattention and State Dependent Stochastic Choice," *working paper*, 2013.
- Cason, Timothy N, Anya C Savikhin, and Roman M Sheremeta, "Behavioral spillovers in coordination games," *European Economic Review*, 2011, 56, 233–245.
- Charles, Thomas, Dean Quinn, Mark Weatherall, Sarah Aldington, Richard Beasley, and Shaun Holt, "An Audiovisual Reminder Function Improves Adherence with Inhaled Corticosteroid Therapy in Asthma," *Journal of Allergy and Clinical Immunology*, 2007, 119 (4), 811–816.
- Charness, Gary and Matthew Rabin, "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 2002, 117 (3), 817–869.
- and Uri Gneezy, "Incentives to Exercise," *Econometrica*, 2009, 77 (3), 909–931.
- Chetty, Raj, "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 2009, 1 451–488.
- Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," NBER Working Paper 13330, 2007.
- Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, 2009, 99 (4) 1145–1177.
- Chen, Yuxin, Sridhar Moorthy, and Z. John Zhang, "Research Note–Price Discrimination After the Purchase: Rebates as State-Dependent Discounts," *Management Science*, 2005, 51 (7), 1131–1140.
- Choi, James, David Laibson, and Brigitte Madria, "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds," *Review of Financial Studies*, 2010, 23 (4) 1405–1432.
- Choi, James, David Laibson, and Brigitte Madrian, "\$100 Bills on the Sidewalk: Suboptimal Investment in 401(k) Plans," *Review of Economics and Statistics*, 2011, 93 (3) 748–763.
- Chun, Marvin M., Julie D. Golomb, and Nicholas B. Turk-Browne, "A Taxonomy of External and Internal Attention," *Annual Review of Psychology*, 2011, 62, 73–101.

- Coffman, Lucas C., "Essays in Experimental Economics," *Doctoral Dissertation*, 2010.
- Conlisk, John, "Optimization Cost," *Journal of Economic Behavior and Organization*, 1988, 9 (3) 213–228.
- Cooper, David J. and E. Glenn Dutcher, "The dynamics of responder behavior in ultimatum games: a meta-study," *Experimental Economics*, 2011, 14, 519–546.
- , Nick Feltovich, Alvin E. Roth, and Rami Zwick, "Relative versus absolute speed of adjustment in strategic environments: responder behavior in ultimatum games," *Experimental Economics*, 2003, 6 (2), 181–207.
- Datta, Somnath and Glen Satten, "Rank-Sum Tests for Clustered Data," *Journal of the American Statistical Association*, 2005, 471 (1), 908–915.
- Davis, Lucas, "Evaluating the Slow Adoption of Energy Efficient Investments: Are Renters Less Likely to Have Energy Efficient Appliances?" in *The Design and Implementation of U.S. Climate Policy*, Edited by Don Fullerton and Catherine Wolfram. University of Chicago Press, 2012.
- Davis, Lucas, and Erich Muehlegger, "Do Americans Consume too Little Natural Gas? An Empirical Test of Marginal Cost Pricing," *RAND Journal of Economics*, 2010, 41 (4) 791–810.
- DeFrank, Jessica T., Barbara K. Rimer, Jennifer M. Gierisch, J. Michael Bowling, David Farrell, and Celette S. Skinner, "Impact of Mailed and Automated Telephone Reminders on Receipt of Repeat Mammograms," *American Journal of Preventive Medicine*, 2009, 36 (6), 259–467.
- DellaVigna, Stefano and Ulrike Malmendier, "Contract Design and Self-Control: Theory and Evidence," *Quarterly Journal of Economics*, 2004, CXIX (2), 353–402.
- Deutsch, Matthias, "Life Cycle Cost Disclosure, Consumer Behavior, and Business Implications: Evidence from an Online Field Experiment," *Journal of Industrial Ecology*, 2010a, 14 (1) 103–120.
- Deutsch, Matthias, "The Effect of Life-Cycle Cost Disclosure on Consumer Behavior: Evidence from a Field Experiment with Cooling Appliances," *Energy Efficiency*, 2010b, 3 303–315.
- DEHWA (Australian Department of the Environment, Water, Heritage, and the Arts), "Regulatory Impact Statement: Proposal to Phase-Out Inefficient Incandescent Light Bulbs, 2008. [http://www.energyrating.gov.au/wp-content/uploads/Energy\\_Rating\\_Documents/Library/Lighting/Incandescent\\_Lamps/200808-09-Phase-out.pdf](http://www.energyrating.gov.au/wp-content/uploads/Energy_Rating_Documents/Library/Lighting/Incandescent_Lamps/200808-09-Phase-out.pdf)
- Diamond, Peter, "Consumption Externalities and Imperfect Corrective Pricing," *Bell Journal of Economics and Management Science*, 1973, 4 (2) 526–538.
- Dismukes, R. Key, "Prospective Memory in Workplace and Everyday Situations," *Current Directions in Psychological Science*, 2012, 21 (4), 215–220.

- DOE (U.S. Department of Energy) Technical Support Document: Impacts on the Nation of the Energy Independence and Security Act of 2007. [http://www1.eere.energy.gov/buildings/appliance\\_standards/pdfs/en\\_masse\\_tsd\\_march\\_2009.pdf](http://www1.eere.energy.gov/buildings/appliance_standards/pdfs/en_masse_tsd_march_2009.pdf)
- DOE (U.S. Department of Energy), “Energy Star CFL Market Profile: Data Trends and Market Insights,” 2010. [http://www.energystar.gov/ia/products/downloads/CFL\\_Market\\_Profile\\_2010.pdf](http://www.energystar.gov/ia/products/downloads/CFL_Market_Profile_2010.pdf),
- DOE (U.S. Department of Energy), “2010 US Lighting Market Characterization,” 2012. <http://apps1.eere.energy.gov/buildings/publications/pdfs/ssl/2010-lmc-final-jan-2012.pdf>
- Dranove, David, and Ginger Zhe Jin, “Quality Disclosure and Certification: Theory and Practice,” *Journal of Economic Literature*, 2010, 48 (4) 935–963.
- Duarte, Fabian, and Justine Hastings, “Fettered Consumers and Sophisticated Firms: Evidence from Mexico’s Privatized Social Security Market,” NBER Working Paper 18582, 2012.
- Dubin, Jeffrey, and Daniel McFadden, “An Econometric Analysis of Residential Electric Appliance Holdings and Consumption,” *Econometrica*, 1984, 52 (2) 345–362.
- Duflo, Esther, and Emmanuel Saez, “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment,” *Quarterly Journal of Economics*, 2003, 118 (3) 815–842.
- Dupas, Pascaline, “Do Teenagers Respond to HIV Risk Information? Evidence From a Field Experiment in Kenya,” *American Economic Journal: Applied Economics*, 2011 3 1–34.
- Dufwenberg, Martin and Georg Kirchsteiger, “A Theory of Sequential Reciprocity,” *Games and Economic Behavior*, 2004, 47 (2), 268–298.
- Edwards, Matthew A., “The Law, Marketing and Behavioral Economics of Consumer Rebates,” *Stanford Journal of Law, Business, and Finance*, 2007, 12 (362), 362–424.
- , “The Rebate “Rip-Off”: New York’s Legislative Responses to Common Consumer Rebate Complaints,” *Pace Law Review*, 2009, 29 (3), 471–478.
- Eliaz, Kfir and Ran Spiegler, “Consideration Sets and Competitive Marketing,” *Theoretical Economics*, 2011, 6, 127–155.
- and —, “On the Strategic Use of Attention Grabbers,” *Review of Economic Studies*, 2011, 78 (1), 235–262.
- and Rani Spiegler, “Reference Dependence and Labor-Market Fluctuations,” Technical Report, CEPR Discussion Papers 2012.
- Energy Star. ENERGY STAR Unit Shipment and Market Penetration Report Calendar Year 2012 Summary, 2013. [http://www.energystar.gov/ia/partners/downloads/unit\\_shipment\\_data/2012\\_USD\\_Summary\\_Report.pdf?e9f6-1224](http://www.energystar.gov/ia/partners/downloads/unit_shipment_data/2012_USD_Summary_Report.pdf?e9f6-1224)

- Ericson, Keith M. Marzilli, "On the Interaction of Memory and Procrastination," 2010, p. working paper.
- , "Forgetting we Forget: Overconfidence and Memory," *Journal of the European Economic Association*, 2011, 9 (1), 43–60.
- Falk, Armin and Florian Zimerrmann, "A Taste for Consistency and Survey Response Behavior," *Cesifo Economic Studies*, 2002 59 (1) 181–193.
- Falk, Armin and Urs Fischbacher, "A theory of reciprocity," *Games and Economic Behavior*, 2006, 54 (2), 293–315.
- , Ernst Fehr, and Christian Zehnder, "Fairness perceptions and reservation wages—the behavioral effects of minimum wage laws," *The Quarterly Journal of Economics*, 2006, 121 (4), 1347–1381.
- Falkinger, Josef, "Attention Economies," *Journal of Economic Theory*, 2007, 133, 266–294.
- , "Limited Attention as a Scarce Resource in Information-Rich Economies," *The Economic Journal*, 2008, 118, 1596–1620.
- Fehr, Ernst and Klaus M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics*, 1999, 114 (3), 817–868.
- , Georg Kirchsteiger, and Arno Riedl, "Does fairness prevent market clearing? An experimental investigation," *The Quarterly Journal of Economics*, 1993, 108 (2), 437–459.
- , Lorenz Goette, and Christian Zehnder, "A Behavioral Account of the Labor Market: The Role of Fairness Concerns," *Annual Review of Economics*, 2009, 1 (1), 355–384.
- , Oliver D. Hart, and Christian Zehnder, "How Do Informal Agreements and Renegotiation Shape Contractual Reference Points?," Working Paper 17545, National Bureau of Economic Research October 2011.
- , Oliver Hart, and Christian Zehnder, "Contracts as Reference Points—Experimental Evidence," *American Economic Review*, 2011, 101 (2), 493–525.
- Finkelstein, Amy, "E-ZTAX: Tax Salience and Tax Rates. *Quarterly Journal of Economics*, 2009 124 (3) 969–1010.
- Figlio, David, and Maurice E. Lucas, "What's in a Grade? School Report Cards and the Housing Market," *American Economic Review*, 2004, 94 (3) 591–604.
- Fischer, Carolyn, Winston Harrington, and Ian Parry, "Do Market Failures Justify Tightening Corporate Average Fuel Economy (CAFE) Standards?" *The Energy Journal*, 2007, 28 (4) 1–30.
- Fischbacher, Urs, "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 2007, 10, 171–178.
- , Christina M. Fong, and Ernst Fehr, "Fairness, errors and the power of competition," *Journal of Economic Behavior & Organization*, 2009, 72 (1), 527–545.



- Fong, Christina, and Erzo F.P. Luttmer, "What Determines Giving to Hurricane Katrina Victims? Experimental Evidence on Racial Group Loyalty," *American Economic Journal: Applied Economics*, 2009, 1 (2) 64–87.
- Formisano, Bob, "2007 Energy Bill – Are They Phasing Out or Making Incandescent Bulbs Illegal?", 2008. [http://homerepair.about.com/od/electricalrepair/ss/2007\\_energybill.htm](http://homerepair.about.com/od/electricalrepair/ss/2007_energybill.htm)
- Frederick, Shane, "Cognitive reflection and decision making," *The Journal of Economic Perspectives*, 2005, 19 (4), 25–42.
- Fudenberg, Drew, David K. Levine, and Zacharias Maniadis, "On the Robustness of Anchoring Effects in WTP and WTA Experiments," *American Economic Journal: Microeconomics*, 2012, 4 (2), 131–145.
- Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality, with Application to Basic Consumer and Equilibrium Theory," *working paper*, 2012.
- and David Laibson, "Bounded Rationality and Directed Cognition," *working paper*, 2005.
- and – , "Shrouded Attributes, Consumer Myopia and Information Supression in Competitive Markets," *Quarterly Journal of Economics*, 2006, 121 (2), 505–540.
- Gerstner, Eitan and James D. Hess, "A Theory of Channel Price Promotions," *The American Economic Review*, 1991, 81 (4), 872–886.
- Gilbert, Ben and Joshua S. Graff Zivin, "Dynamic Salience with Intermittent Billing: Evidence from Smart Electricity Meters," *NBER Working Paper No. 19510*, 2013.
- Gilpatric, Scott M., "Slippage in Rebate Programs and Present-Biased Preferences," *Marketing Science*, 2009, 28 (2), 229–238.
- Gillingham, Kenneth, Matthew Harding, and David Rapson, "Split Incentives and Household Energy Consumption." *Energy Journal*, 2012, 33 (2) 37–62.
- Gillingham, Kenneth, and Karen Palmer, "Bridging the Energy Efficiency Gap," Resources for the Future Discussion Paper 13–02, 2013.
- Goldberg, Pinelopi, "The Effects of the Corporate Average Fuel Economy Standards in the US," *Journal of Industrial Economics*, 1998, 46 1–33.
- Greenstone, Michael, Paul Oyer, and Annette Vissing-Jorgensen, "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments," 2006, *Quarterly Journal of Economics*, 121 (2) 399–460.
- Greiner, Ben, "The Online Recruitment System ORSEE: A Guide for the Organization of Experiments in Economics," Discussion Papers on Strategic Interaction 2003–10, Max Planck Institute of Economics, Strategic Interaction Group 2004.
- Grimm, Veronika and Friederike Mengel, "An experiment on learning in a multiple games environment," *Journal of Economic Theory*, 2012, 147, 2220–2259.

- Grubb, Michael D., "Consumer Inattention and Bill-Shock Regulation," *Working Paper*, 2012.
- Gruber, Jonathan, and Botond Köszegi, "Tax Incidence when Individuals are Time-Inconsistent: The Case of Cigarette Excise Taxes," *Journal of Public Economics*, 2004, 88 1959–1987.
- Gul, Faruk, and Wolfgang Pesendorfer, "Harmful Addiction," *Review of Economic Studies*, 2007 74 (1) 147–172.
- Güth, Werner, Rolf Schmittberger, and Bernd Schwarze, "An experimental analysis of ultimatum bargaining," *Journal of Economic Behavior & Organization*, 1982, 3 (4), 367–388.
- Haapala, Irja, Noel C. Barengo, Simon Biggs, Leena Surakka, and Pirjo Manninen, "Weight Loss by Mobile Phone: A 1-year Effectiveness Study," *Public Health Nutrition*, 2012, 12, 2382–2391.
- Hall, Robert E, "Employment fluctuations with equilibrium wage stickiness," *American economic review*, March 2005, 95 (1), 50–65.
- Hart, Oliver, "Hold-up, Asset Ownership, and Reference Points," *The Quarterly Journal of Economics*, 2009, 124 (1), 267–200.
- and Bengt Holmstrom, "A Theory of Firm Scope," *The Quarterly Journal of Economics*, 2010, 125 (2), 483–513.
- and John Moore, "Contracts as Reference Points," *The Quarterly Journal of Economics*, 2008, 123 (1), 1119–11158.
- Hassett, Kevin, and Gilbert Metcalf, "Energy Tax Credits and Residential Conservation Investment: Evidence from Panel Data," *Journal of Public Economics*, 57 201—217.
- Hastings, Justine, and Jeffrey Weinstein, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics*, 2008, 123 (4) 1373–1414.
- Hausman, Jerry, "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables," *Bell Journal of Economics*, 1979, 10 (1) 33–54.
- Hay, Janine F. and Larry L. Jacoby, "Separating Habit and Recollection: Memory Slips, Process Dissociations, and Probability Matching," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1996, 22 (6), 1323–1335.
- Hefti, Andreas M., "Attention and Competition," *working paper*, 2013.
- Heidhues, Paul, Botond Köszegi, and Takeshi Murooka, "Deception and Consumer Protection in Competitive Markets," in "Anthology on Pros and Cons of Consumer Protection" 2011.
- , —, and —, "Exploitative Innovation," *working paper*, 2012.
- , —, and —, "Inferior Products and Profitable Deception," *working paper*, 2012.

Heiss, Florian, Daniel McFadden, and Joachim Winter, "Mind the Gap! Consumer Perceptions and Choices of Medicare Part D Prescription Drug Plans," NBER Working Paper 13627, 2007

Henderson, Conor M., Joshua T. Beck, and Robert W. Palmatier, "Review of the Theoretical Underpinnings of Loyalty Programs," *Journal of Consumer Psychology*, 2011, 21, 256–276.

Herberich, David, John List, and Michael Price, "How Many Economists Does it Take to Change a Light Bulb? A Natural Field Experiment on Technology Adoption," Working Paper, University of Chicago, 2011.

Herrnstein, R. J., George Loewenstein, Drazen Prelec, and William Vaughan, Jr., "Utility Maximization and Melioration: Internalities in Individual Choice," *Journal of Behavioral Decision Making*, 1993 6 149–185.

Heutel, Garth, "Optimal Policy Instruments for Externality-Producing Durable Goods under Time Inconsistency," Working Paper, UNC Greensboro, 2011.

Holman, Jeff and Farhan Zaidi, "The Economics of Prospective Memory," *working paper*, 2010.

Hossain, Tanjim, and John Morgan, "...Plus Shipping and Handling: Revenue (Non)Equivalence in Field Experiments on eBay." *Advances in Economic Analysis and Policy*, 6, 2006.

Houde, Sebastien (2012), "How Consumers Respond to Certification: A Welfare Analysis of the Energy Star Program," Working Paper, Stanford University, 2012.

Interagency Working Group on Social Cost of Carbon, United States Government, "Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866, 2013. [http://www.whitehouse.gov/sites/default/files/omb/inforeg/social\\_cost\\_of\\_carbon\\_for\\_ria\\_2013\\_update](http://www.whitehouse.gov/sites/default/files/omb/inforeg/social_cost_of_carbon_for_ria_2013_update).

Insel, Kathleen C., Gilles O. Einstein, Daniel G. Morrow, and Joseph T. Hepworth, "A Multifaceted Prospective Memory Intervention to Improve Medication Adherence: Design of a Randomized Control Trial," *Contemporary Clinical Trials*, 2013, 34, 45–52.

Jacoby, Larry L., James A. Debner, and Janine F. Hay, "Proactive Interference, Accessibility Bias, and Process Dissociations: Valid Subjective Reports of Memory," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2001, 27 (3), 686–700.

Jaffe, Adam, and Robert Stavins, "The Energy Paradox and the Diffusion of Conservation Technology," *Resource and Energy Economics*, 1994 16 91–122.

Janakiraman, Narayan and Lisa Ordóñez, "Effect of Effort and Deadlines on Consumer Product Returns," *Journal of Consumer Psychology*, 2012, 22, 260–271.

Jensen, Robert, "The Perceived Returns to Education and the Demand for Schooling." *Quarterly Journal of Economics*, 2010 125 (2) 515–548.

- Jessoe, Katrina and David Rapson, "Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use," *American Economic Review*, forthcoming.
- Jin, Ginger Zhe, and Alan Sorensen, "Information and Consumer Choice: The Value of Publicized Health Plan Ratings." *Journal of Health Economics*, 2006, 25 (2) 248–275.
- Kagel, John H, Chung Kim, and Donald Moser, "Fairness in ultimatum games with asymmetric information and asymmetric payoffs," *Games and Economic Behavior*, 1996, 13 (1), 100–110.
- Kahneman, Daniel and Amos Tversky, *Choices, Values, and Frames*, 1st ed., New York: Cambridge University Press, 2000.
- , Jack L Knetsch, and Richard Thaler, "Fairness as a constraint on profit seeking: Entitlements in the market," *The American economic review*, September 1986, 76 (4), 728–741.
- Kallbekken, Steffen, Hakon Saelen, and Erlend Hermansen, " Bridging the Energy Efficiency Gap: A Field Experiment on Lifetime Energy Costs and Household Appliances," *Journal of Consumer Policy*, 2013, 36 1—16.
- Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman, "Getting to the Top of Mind: How Reminders Increase Saving," *working paper*, 2012.
- , Melanie Morten, and Jonathan Zinman, "A personal touch: Text messaging for loan repayment," 2012, p. working paper.
- Kaur, Supreet, "Nominal Wage Rigidity in Village Labor Markets," *Working Paper*, 2012.
- Kling, Jeffrey, Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian Wrobel, "Comparison Friction: Experimental Evidence from Medicare Drug Plans," *Quarterly Journal of Economics*, 2012, 127 (1) 199–235.
- Kőszegi, Botond and Matthew Rabin, "A Model of Reference-Dependent Preferences.," *Quarterly Journal of Economics*, 2006, 121 (4), 1133–65.
- and —, "Reference-Dependent Risk Attitudes," *American Economic Review*, 2007, 99 (3), 909–936.
- and —, "Reference-Dependent Consumption Plans," *American Economic Review*, 2009, 99 (3), 909–936.
- Kőszegi, Botond and Adam Szeidl, "A Model of Focusing in Economic Choice," *Quarterly Journal of Economics*, 2013, 128 (1), 53–107.
- Krishna, Santosh, Suzanne Austin Boren, and E. Andrew Balas, "Healthcare via Cell Phones: A Systematic Review," *Telemedicine and e-Health*, 2009, 15 (3), 231–240.
- Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, 112 (2), 443–477.
- , "A Cue-Theory of Consumption," *Quarterly Journal of Economics*, 2001, 116 (1), 81–119.

- Letzler, Robert and Joshua Tasoff, "Everyone Believes in Redemption: Nudges and Overoptimism in Costly Task Completion," *working paper*, 2013.
- Levine, David K., "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, 1998, 1, 593–622.
- Levitt, Steve, and John List, "Field Experiments in Economics: The Past, the Present, and the Future." *European Economic Review*, 2009, 53 1—18.
- Lichtenstein, Sarah and Paul Slovic, *The construction of preference*, 1st ed., New York: Cambridge University Press, 2006.
- List, John A., Zacharias Maniadis, and Fabio Tufano, "One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects," *American Economic Review*, *forthcoming*, 2013.
- List, John and Todd Cherry, "Learning to accept in ultimatum games: evidence from an experimental design that generates low offers," *Experimental Economics*, 2000, 3 (1), 11–29.
- Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics*, 2003, 118 (4), 1209–1248.
- Lu, Qiang and Sridhar Moorthy, "Coupons Versus Rebates," *Marketing Science*, 2007, 26 (1), 67–82.
- Luo, Renfu, Yaojian Shi, Linxiu Zhang, Huiping Zhang, Grant Miller, Alexis Medina, and Scott Rozelle (2012), "The Limits of Health and Nutrition Education: Evidence from Three Randomized-Controlled Trials in Rural China," *CESifo Economic Studies*, 2012, 58 (2) 385–404.
- Lynch, John G. and Gal Zauberan, "When do you Want it? Time, Decisions, and Public Policy," *American Marketing Association*, 2006, 25 (1), 67–78.
- MacDonell, Karen, Sylvie Naar-King, Heather Huszti, and Marvin Belzer, "Barriers to Medication Adherence in Behaviorally and Perinatally Infected Youth Living with HIV," *AIDS Behavior*, 2013, 17 (1), 86–93.
- Martin, Benjamin A., Noelle L. Brown, and Jason L. Hicks, "Ongoing Task Delays Affect Prospective Memory More Powerfully Than Filler Task Delays," *Canadian Journal of Experimental Psychology*, 2011, 65 (1), 48–56.
- McBride, Dawn M., Jaclyn K. Beckner, and Drew H. Abney, "Effects of Delay of Prospective Memory Cues in an Ongoing Task on Prospective Memory Task Performance," *Memory & Cognition*, 2011, 39 (7), 1222–1231.
- , Jennifer H. Coane, Jason Drwal, and Stephanie-Ann M. LaRose, "Differential Effects of Delay on Time-based Prospective Memory," *Aging, Neuropsychology, and Cognition*, in press.
- Metcalf, Gilbert, and Kevin Hassett, "Measuring the Energy Savings from Home Improvement Investments: Evidence from Monthly Billing Data," *Review of Economics and Statistics*, 81 (3) 517–528.

- Mills, Bradford, and Joachim Schleich, "Why Don't Households See the Light? Explaining the Diffusion of Compact Fluorescent Lamps," *Resource and Energy Economics*, 2010, 32 363–378.
- Mitra, Anusree and John G. Lynch., "Toward a Reconciliation of Market Power and Information Theories of Advertising Effects on Price Elasticity," *Journal of Consumer Research*, 1995, 21 (4), 644–659.
- Mullainathan, Sendhil, "A Memory-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2002, 117 (3), 735–774.
- Mullainathan, Sendhil, Joshua Schwartzstein, and William J. Congdon, "A Reduced-Form Approach to Behavioral Public Finance," *Annual Review of Economics*, 2012, 4 17.1—17.30.
- Narasimhan, Chakrvarthi, "A Price Discrimination Theory of Coupons," *Marketing Science*, 1984, 3 (2), 128–147.
- Nedungadi, Prakash, "Recall and Consumer Consideration Sets: Influencing Choice without Altering Brand Evaluations," *Journal of Consumer Research*, 1990, 17 (3), 263–276.
- NHTSA (National Highway Traffic Safety Administration), "Final Regulatory Impact Analysis: Corporate Average Fuel Economy for MY 2012–MY 2016 Passenger Cars and Light Trucks," Office of Regulatory Analysis and Evaluation, National Center for Statistics and Analysis, 2010.
- Newell, Richard, and Juha Siikamaki, "Nudging Energy Efficiency Behavior: The Role of Information Labels," Resources for the Future Discussion Paper 13–17, 2013.
- NRDC (Natural Resources Defense Council), "Better Light Bulbs Equal Consumer Savings in Every State, 2011. [www.nrdc.org/policy](http://www.nrdc.org/policy)
- O'Donoghue, Ted and Matthew Rabin, "Doing it Now or Later," *American Economic Review*, 1999, 89 (1), 103–124.
- O'Donoghue, Edward, and Matthew Rabin, "Optimal Sin Taxes," *Journal of Public Economics*, 2006, 90 1825–1849.
- Oreopoulos, Philip, Till von Wachter, and Andrew Heisz, "The Short-and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, 2012, 4 (1), 1–29.
- Parry, Ian, David Evans, and Wallace Oates, "Are Energy Efficiency Standards Justified?" Resources for the Future Discussion Paper 10–59, 2010.
- Patrick, Kevin, Fred Raab, Marc A. Adams, Lindsay Dillon, Marian Zabinski, Cheryl L. Rock, William G. Griswold, and Gregory J. Norman, "A Text Message-Based Intervention for Weight Loss: Randomized Controlled Trial," *Journal of Medical Internet Research*, 2009, 11.
- Peng, Lin and Wei Xiong, "Investor attention, overconfidence and category learning," *Journal of Financial Economics*, 2006, 80, 563–602.

- Peysakhovich, Alexander and David G. Rand, "Habits of Virtue: Creating Norms of Cooperation and Defection in the Laboratory," *working paper*, 2013.
- Pope, Devin, "Reacting to Rankings: Evidence from 'America's Best Hospitals.'" *Journal of Health Economics*, 2009, 28 (6) 1154–1165.
- Poquette, Amelia J., David J. Moore, Ben Gouaux, Erin E. Morgan, Igor Grant, and Steven P. Woods, "Prospective Memory and Antiretroviral Medication Non-Adherence in HIV: An Analysis of Ongoing Task Delay Length Using the Memory for Intentions Screening Test," *Journal of the International Neuropsychological Society*, 2013, 19, 155–161.
- Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 1993, 83 (5), 1281–1302.
- Rabin, Matthew, and Georg Weizsacker, "Narrow Bracketing and Dominated Choices." *American Economic Review*, 2009, 99 (4) 1508–1543.
- Rankin, Catharine H., ..., and Richard F. Thompson, "Habituation Revisited: An Updated and Revised Description of the Behavioral Characteristics of Habituation," *Neurobiology of Learning and Memory*, 2009, 92, 135–138.
- Reis, Ricardo, "Inattentive Consumers," *Journal of Monetary Economics*, 2006, 53, 1761–1800.
- Roth, Alvin E. and Francoise Schoumaker, "Expectations and Reputations in Bargaining: An Experimental Study," *The American Economic Review*, 1983, 73 (3), 362–372.
- and Ido Erev, "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 1995, 8, 164–212.
- and – , "Predicting How people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria.," *American Economic Review*, 1998, 88 (4), 848–881.
- Roth, Alvin E, Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir, "Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study," *The American Economic Review*, December 1991, 81 (5), 1068–1095.
- Sallee, James, "Rational Inattention and Energy Efficiency," Working Paper, University of Chicago, 2013.
- Sallee, James, Sarah West, and Wei Fan, "Consumer Valuation of Fuel Economy: A Microdata Approach," Working Paper, University of Chicago, 2009. Available from <http://www.ntanet.org/images/stories/pdf/proceedings/09/032.pdf>.
- Sanstad, Alan, and Richard Howarth, "'Normal' Markets, Market Imperfections, and Energy Efficiency," 1994 *Energy Policy*, 22 (10) 811–818.
- Scanlon, Dennis, Michael Chernew, Catherine McLaughlin, and Gary Solon, "The Impact of Health Plan Report Cards on Managed Care Enrollment," *Journal of Health Economics*, 2002, 21 (1) 19–41.

- Schmidt, Klaus M. and Fabian Herweg, "A Theory of Ex Post Inefficient Renegotiation," *GESY Discussion Paper No. 390*, 2012.
- Schwartzstein, Josh, "Selective Attention and Learning," *working paper*, 2012.
- Sellen, A.J., G. Louie, J.E. Harris, and A.J. Wilkins, "What Brings Intentions to Mind?," *Memory*, 1997, 5 (4), 483–507.
- Shafir, Eldar and Amos Tversky, "Choice Under Conflict: The Dynamics of Deferred Decision," *Psychological Science*, 1992, 3, 358–361.
- Shimer, Robert, "The cyclical behavior of equilibrium unemployment and vacancies," *American economic review*, March 2005, 95 (1), 25–49.
- Shu, Suzanne B. and Ayelet Gneezy, "Procrastination of Enjoyable Experiences," *Journal of Marketing Research*, 2010, 47, 933–944.
- Silk, Tim and Chris Janiszewski, "Managing Mail-in Rebate Promotions," 2008.
- Silk, Timothy, "Examining Purchase and non-Redemption of Mail-in Rebates: The Impact of Offer Variables on Consumers' Subjective and Objective Probability of Redeeming," 2004.
- Simonsohn, Uri and George Loewenstein, "Mistake# 37: The Effect of Previously Encountered Prices on Current Housing Demand\*," *The Economic Journal*, 2006, 116 (508), 175–199.
- Simonson, Itamar and Amos Tversky, "Choice in context: tradeoff contrast and extremeness aversion.," *Journal of Marketing Research*, 1992.
- Sims, Christopher A., "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, 50 (3), 665–690.
- Slonim, Robert and Alvin E. Roth, "Learning in high stakes ultimatum games: An experiment in the Slovak Republic," *Econometrica*, 1998, 66 (3), 569–596.
- Snyder, Mark, "Self-Monitoring of Expressive Behavior," *Journal of Personality and Social Psychology*, 1974, 30 526–537.
- Sovern, Jeff, "Toward a New Model of Consumer Protection: The Problem of Inflated Transaction Costs," *William and Mary Law Review*, 2006, 47 (5), 1635–1709.
- Spiegler, Ran, "Competitive Framing," *American Economic Journal: Microeconomics*, forthcoming.
- Stigler, George J., "The Economics of Information," *Journal of Political Economy*, 1961, 69 (3), 213–225.
- Strotz, Robert, "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 1956 23 165–180.
- Sunstein, Cass and Richard Thaler, "Libertarian Paternalism is not an Oxymoron," *University of Chicago Law Review*, 2003, 70 (4) 1159–1202.



- Sylvania, "5th Annual Sylvania Socket Survey, 2012. <https://www.sylvania.com/en-us/tools-and-resources/surveys/Pages/socket-survey.aspx>
- Szilagyi, Peter C., Clayton Bordley, Julie C. Vann, Ann Chelminski, Ronald M. Kraus, Peter A. Margolis, and Lance E. Rodewald, "Effect of Patient Reminder/Recall Interventions on Immunization Rates," *Journal of the American Medical Association*, 2000, 284 (14), 1820–1827.
- Tobias, Robert, "Changing Behavior by Memory Aids: A Social Psychological Model of Prospective Memory and Habit Development Tested With Dynamic Field Data," *Psychological Review*, 2009, 116, 408–438.
- Toledo, Chantal, "Do Environmental Messages Work on the Poor? Experimental Evidence from Brazilian Favelas," Working Paper, University of California at Berkeley, 2013.
- Vervloet, Marcia, Annemiek J. Linn, Julia C. M. van Weert, Dinny H. de Bakker, Marcel L. Bouvy, and Liset van Dijk, "The Effectiveness of Interventions Using Electronic Reminders to Improve Adherence to Chronic Medication: A systematic Review of the Literature," *Journal of the American Medical Association*, 2012, 19, 696–704.
- Villas-Boas, J. Miguel, "Predicting Advertising Pulsing Policies in an Oligopoly: A Model and Empirical Test," *Marketing Science*, 1993, 12 (1), 88–102.
- Vyankandondera, Joseph, Kirstin Mitchell, Brenda Asiimwe-Kateera, Kimberly Boer, Philippe Mutwa, Jean-Paul Balinda, Masja van Straten, Peter Reiss, and Janneke van de Wijgert, "Antiretroviral therapy drug adherence in Rwanda: Perspectives from patients and healthcare workers using a mixed-methods approach," *AIDS care*, 2013, 25 (12), 1504–1512.
- Ward, David, Christopher Clark, Kimberly Jensen, Steven Yen, and Clifford Russell, "Factors Influencing Willingness-to-Pay for the Energy Star Label," *Energy Policy*, 2011, 39 (3) 1450–1458.
- Woodford, Michael, "Inattentive Valuation and Reference-dependent Choice," 2012.
- Woodruff, Mandi, "Genius PhD Student Paid His Tuition With Mail-In-Rebates," September 2012. <http://www.businessinsider.com/a-genius-phd-student-paid-his-tuition-almost-entirely-with-mail-in-rebates-2012-9>.
- Yantis, Steven, "Control of Visual Attention," in H. Pashler, ed., *Attention*, The Guilford Press, 1998, pp. 223–156.
- Zogg, Jennifer B., Steven P. Woods, John A. Saucedo, John S. Wiebe, and Jane M. Simoni, "The Role of Prospective Memory in Medication Adherence: A Review of an Emerging Literature," *Journal of Behavioral Medicine*, 2012, 35 (1), 47–62.
- Zwick, Rami and Vincent Mak, "Gaming with Fairness: Some Conjectures on Behavior in Alternating Offer Bargaining Experiments," in Gary Bolton and Rachel Croson, eds., *Conflict Resolution Handbook*, The Oxford University Press, 2012, pp. 91–107.
- and Xiao-Ping Chen, "What Price Fairness? A Bargaining Study," *Management Science*, 1999, 45 (6), 804–823.

# Appendix A

## Appendix to Chapter 1

### A.1 Extensions and Additional Examples

#### A.1.1 Infinite Horizon

Let  $\delta < 1$  be the DM's discount factor, and suppose now that  $T$  is possibly infinite. For simplicity, I will assume that period  $t$  flow payoffs are independent of the history  $t$ . Let  $H_t$  be the set of all possible period  $t$  histories, and set  $\mathbf{H} \equiv \bigcup H_t$ . Let  $\mathbf{x}^s : \mathbf{H} \times \mathbb{R} \rightarrow \{d, a\}$  denote a sophisticated DM's plan, which maps each pair  $(h_t, \xi_t)$  to an action in  $A(h_t)$ , *conditional on the DM being attentive*. Let  $z_t = (x_t, \xi_t)$  denote the period  $t$  outcome, and let  $u_t(z_t)$  denote the period  $t$  flow payoff corresponding to that outcome. Finally, let  $\mathcal{F}(\mathbf{x}^s)$  be the distribution over the possible outcomes  $z = (z_0, z_1, \dots)$  induced by strategy  $\mathbf{x}^s$ . Note that the product measure  $\mathcal{F}(\mathbf{x}^s)$  takes into account how the likelihood of being attentive evolves over time as a function of past events.

The sophisticated DM chooses the strategy  $\mathbf{x}^s$  that maximizes

$$\sum_{t=0}^{\infty} \delta^t u(z) d\mathcal{F}^s(\mathbf{x}^s).$$

The naive DM chooses the strategy  $\mathbf{x}^n$  that corresponds to the strategy of a perfectly attentive DM. In particular, let  $\mathcal{F}^{\text{pa}}(\mathbf{x}^n)$  denote the product measure over outcomes that would be induced if a perfectly attentive DM used strategy  $\mathbf{x}^n$ . The naive DM chooses the

strategy  $\mathbf{x}^n$  that maximizes

$$\sum_{t=0}^{\infty} \delta^t u(z) d\mathcal{F}^{\text{pa}}(\mathbf{x}^n).$$

His utility, however, is given by

$$\sum_{t=0}^{\infty} \delta^t u(z) d\mathcal{F}^n(\mathbf{x}^n)$$

where  $\mathcal{F}^n(\mathbf{x}^n)$  is the actual product measure over a naive DMs' outcomes that  $\mathbf{x}^n$  induces.

### A.1.2 Endogenous Cue Generation

#### Does setting cues eliminate inattention for sophisticates?

The possibility to manipulate cues does not eliminate the DM's inattention problem for a number of reasons. As discussed in Section 1.2.4, cues are imperfect. Second, cues can be costly. While setting a single electronic calendar reminder is cheap, purchasing an electronic pill bottle with an array of audio and visual reminders is more costly. Even electronic calendar reminders, however, can be expensive. Because one calendar reminder is highly imperfect in the sense that it can be forgotten minutes later, ensuring perfect attentiveness over the course of even a 1 week period might require hundreds, if not thousands of reminders—which carries a high nuisance cost. Third, an inattentive DM will not only be inattentive about the primary behavior: he will also be inattentive about setting additional cues. Especially when combined with naivete, these issues substantially reduce the possibility that a DM might eliminate his inattention through reminder technologies.

#### Formal model

I now explore an augmented model in which the DM can set cues in period 0. I begin by focusing on sophisticated DMs. As in section 1.4, suppose that the DM starts out with initial attention probabilities  $\gamma_t^0(\alpha, x)$  but can modify them to  $\gamma_t^{\kappa_t}(\alpha, x)$  at cost  $C(\kappa)$ , where  $\kappa = (\kappa_1, \dots, \kappa_T)$ . For these DMs, let  $V_0^s(\kappa)$  be the period 0 utility as a function of modified attention probabilities, with the  $V_t^s$  still defined as in Section 1.2.3. Different from Section 1.2.3, I now allow the vector  $\kappa$  to be set by the DM in period 0. The period 0 optimization is

now

$$\max_{\kappa \in [0,1]^T} \{V_0^s(\kappa) - C(\kappa)\}, \quad (\text{A.1})$$

Notice that all of the analysis in the paper is just a subgame of this more general framework. Notice also that this general framework still incorporates the idea that some cues are generated exogenously, either by an interested party or as incidental events such as conversations with others.

A basic prediction of this more general framework is that as stakes become larger, the DM invests more in increasing his attention. This would not modify most of the results, however. I now walk through each of the paper's results to check their robustness. To consider analogs to the results about creating additional cues, as in several propositions in the paper, assume that the additions are applied to the initial cue distributions  $H_t$ , but that the DM's cost function  $C(\kappa)$  does not change.

An assumption I will rely on throughout the analysis is that the DM cannot guarantee perfect attentiveness:

**Assumption B** For  $\kappa = (\kappa_1, \dots, \kappa_T)$ ,  $\lim_{\kappa_t \rightarrow 1} C(\kappa) = \infty$  for each  $t$ .

Prop 1 Holds under Assumption B (which is needed to guarantee that there is scope for "external" cues to affect inattention). The logic is slightly more involved, though, because increasing  $b_{t'}$  will now increase both investments in reminders, as well as investments through behavioral rehearsal.

Prop 2 Part 4 goes through exactly. The higher the period  $t'$  cues, the less the DM invests in future attentiveness both through rehearsal and through reminder technologies. Thus the likelihood of him taking an action in period  $t < t'$  decreases. The logic of 2b goes through under assumption B: A sufficiently large increase in period  $t$  cues cannot be crowded out by lack of investment through rehearsal or own reminder technologies. Finally, whether or not 3a goes through under assumption B depends on the curvature of  $C$  and on the function  $g(x, \alpha, \sigma)$ . It is possible that increasing period 1 cues crowds out investment in future cues more than one-for-one. However, in the simple model of

multiplicative cues introduced in (1.1), investment in cues will be crowded out less than one-for-one when  $C$  is continuous and convex. This ensures that increasing  $H_1$  always increases the likelihood of attentiveness in all future periods. Part 5 will hold under these same assumptions.

Prop 3 Whether or not part 2 goes through under assumption B depends on the curvature of  $C$  on the function  $g(x, \alpha, \sigma)$ . As argued above, in the simple model of multiplicative cues introduced in (1.1), investment in cues will be crowded out less than one-for-one when  $C$  is continuous and convex. This ensures that increasing  $H_1$  always increases the likelihood of attentiveness in all future periods.

Prop 4 The logic goes through under assumption B. Making the DM close to attentive as in part (b) makes him close to non-responsive to rehearsal effects (or investments in reminder technologies) from prior periods.

Prop 5 Part (2) of Proposition 5 will hold under assumption B. As discussed in Section 1.2.5, no cue is perfect because there is always a chance that it will be ignored or missed. Part 3 of Proposition 5 may not necessarily hold for very high benefit tasks, since the DM will set many reminders for those tasks. However, this claim should still hold for low benefit tasks for which the DM will not invest as much in reminders.

Prop 7 Parts 1 and 2 clearly hold verbatim, since the more attentive the DM is, the more advantageous the longer deadline. Part 3 will also hold but possibly through a different mechanism: the more important the task, the more attentive the DM becomes to it through endogenous cue setting, thus approaching the behavior of the perfectly attentive DM.

Prop 8 Holds verbatim, since the probability of being attentive is still bounded away from 0.

Prop 9 The analog to part 2 will still hold, in the sense that if the  $\gamma_t^0(\alpha, x)$  are sufficiently high to make the DM close to attentive, then longer deadlines should generate higher completion rates.

Props 13,14 The major new question that the more general framework raises is the extent to which third-party cue-provision crowds out decision makers' personal cue provision. As long as crowd-out is not one-for-one, however, the impact of third-party cues should not go to zero in the long run. The DM's own personal provision may also be non-monotonic in  $\ell$ , thus dampening the extent to which the third party's optimal cue intensity may be non-monotonic in  $\ell$ .

### A.1.3 Modeling Partial Naivete

At the most general level, prediction mistakes can be incorporated by supposing that the DM's forecasts are given by  $\hat{g}(\alpha, x, \sigma)$  that satisfies assumptions A1-A4, but that doesn't necessarily correspond to the true  $g$ . A simple and parametric way of modeling overconfidence is to set

$$\hat{g}(\alpha, x, \sigma) = \chi_o + (1 - \chi_o)g(\alpha, x, \sigma) \quad (\text{A.2})$$

Then  $\chi_o = 0$  corresponds to a fully sophisticated DM,  $\chi_o = 1$  corresponds to a fully naive DM, and  $\chi \in (0, 1)$  corresponds to a partially overconfident DM.

The DM may be naive in more nuanced ways, however. For example, the DM may recognize that he can be inattentive in the future, but be fully naive about the role that behavioral rehearsal plays in increasing accessibility. This can be captured by setting  $\hat{g}(\alpha, a, \sigma) = \hat{g}(\alpha, d, \sigma) = \chi_r g(\alpha, d, \sigma) + (1 - \chi_r)g(\alpha, a, \sigma)$ . In this formulation of naivete, the DM will always overestimate his future attentiveness for one time tasks such as the ones studied in Section 1.4. In the repeated action environments studied in Section 1.3, this formulation will lead the DM to be completely ignorant of the effect that past behavior will have on his future behavior. This ignorance of the behavioral rehearsal effect will sometimes lead the DM to overestimate and sometimes underestimate the probability of choosing  $x_t = a$  in the future.

Even more generally, the partial overconfidence in equation (A.2) can be combined with the naivete about rehearsal described in the previous paragraph.

### A.1.4 Response elasticities for section 1.3.1

Let  $D_t^i(b_t) = e + (1 - Pr^{H^i}(a_t = 1))$  be the expected total consumption of energy in period  $t$ , conditional on sequence of cue distributions  $\mathbf{H}^i$ . Here,  $e \geq 0$  is a baseline electricity use from other decisions. Then  $\frac{\partial}{\partial b_t} D_t^i(b_t) = -\frac{\partial}{\partial b_t} Pr^{H^i}(x_t = a)$ .

Let  $s_t$  be the naive or sophisticated DM's threshold rule in period  $t$ , so that  $x_t = a$  if and only if  $\xi_t \geq s_t$ . Then for  $i = 1, 2$ ,

$$\frac{\frac{\partial Pr^{H^i}(x_t = a)}{\partial b_t}}{Pr^{H^i}(x_t = a)} = \frac{\partial}{\partial b_t} (1 - F(s_t))$$

is constant in  $i$ . To simplify notation, set  $k_t = \frac{\partial}{\partial b_t} (1 - F(s_t))$ . Then

$$\begin{aligned} -\frac{\frac{\partial}{\partial b_t} D_t^1(b_t)}{D_t^1(b_t)} &< -\frac{\frac{\partial}{\partial b_t} D_t^2(b_t)}{D_t^2(b_t)} \\ \Leftrightarrow -\frac{k_t Pr^{H^1}(x_t = a)}{e + (1 - Pr^{H^1}(x_t = a))} &< -\frac{k_t Pr^{H^2}(x_t = a)}{e + (1 - Pr^{H^2}(x_t = a))} \\ \Leftrightarrow -Pr^{H^1}(x_t = a) (e + 1 - Pr^{H^2}(x_t = a)) &< -Pr^{H^2}(x_t = a) (e + 1 - Pr^{H^1}(x_t = a)) \\ \Leftrightarrow -Pr^{H^1}(x_t = a)(e + 1) &< -Pr^{H^2}(x_t = a)(e + 1) \\ \Leftrightarrow Pr^{H^1}(x_t = a) &> Pr^{H^2}(x_t = a) \end{aligned}$$

Thus, when cues increase the probability of  $x_t = a$ , they also increase the demand response elasticity.

### A.1.5 Microfoundations for the price floor assumption

An arbitrageur who derives no intrinsic value from the product gets a total payoff of  $\hat{v}_a - p + r$  from the offer, where  $\hat{v}_a$  corresponds to the scrap value of the product, net of potential inconvenience costs of going through with the deal. I assume that  $\hat{v}_a < c$ . In contrast to the inattentive consumers, an arbitrageur is perfectly attentive, and always mails in the form correctly. Thus, conditional on purchasing a product with a rebate, he obtains the rebate with probability 1. This is in contrast to naive consumers who derive utility  $v$  from one unit of the product but derive 0 utility from further purchases.

**Proposition 23.** *Conditional on any deadline  $T$ , the profit maximizing choice of  $p_T$  and  $r_T$  is such that  $p_T - r_T \geq \hat{v}_a$ .*

*Proof.* If  $p_T - r_T < \hat{v}_a$  then arbitrageurs will have infinite demand for the product. However, the profit from selling to each arbitrageur will be  $p_T - r_T - c < \hat{v}_a - c < 0$ .  $\square$

Note that it is possible there are some arbitrageurs who are intrinsically interested in the product, and derive utility  $v - p + r > \hat{v}_a - p + r$  from the first unit. The analysis in the body of the paper is a limit case of an economy in which the fraction  $\epsilon$  of such arbitrageurs approaches 0.

## A.2 Daily Action Experiment: Supplementary Material

### A.2.1 Theoretical Extensions

To formally model the decision environment in experiment 2, suppose now that the payoff to choosing  $a_t = 1$  is  $b_t + \xi_t + \zeta_t$ , where  $b_t$  and  $\xi_t \sim F$  are as in Section 1.3. The new source of variation is the correlated shock  $\zeta_t$ . Specifically, there are  $t_1$  and  $t_2$  such that

- $\zeta_t = \bar{z}$  for all  $t \leq t_1$  and  $t > t_2$
- With probability  $p$ ,  $\zeta_t = \underline{\zeta}$  for all  $t_1 < t \leq t_2$ , while with probability  $1 - p$ ,  $\zeta_t = \bar{\zeta}$  for all  $t_1 < t \leq t_2$

**Proposition 24.** *For all  $t > t_2$ ,  $t_1 < \tau \leq t_2$ :  $Pr^s(a_t = 1)$  and  $Pr^n(a_t = 1)$  are increasing in the realization  $\zeta_\tau$ .*

**Proposition 25.** *Consider two different sequences of cue distributions  $\mathbf{H}^1 = (H_1^1, \dots, H_T^1)$  and  $\mathbf{H}^2 = (H_1^2, \dots, H_T^2)$  such that  $H_\tau^1 = H_\tau^2$  for  $\tau \leq t_2$ , but such that  $H_\tau^2 \geq H_\tau^1$  for  $\tau > t_2$ , with strict equality for at least one such  $\tau$ .*

1. For all  $t > t_2$ ,  $t_1 < \tau \leq t_2$ :  $Pr^{n,H^2}(a_t = 1) - Pr^n(a_t = 1)$  is decreasing in  $\zeta_\tau$ .
2. If  $\underline{\zeta}$  is sufficiently low then for all  $t > t_2$  and  $t_1 < \tau \leq t_2$ ,  $Pr^{s,H^2}(a_t = 1) - Pr^{s,H^1}(a_t = 1)$  is decreasing in  $\zeta_\tau$ .



3. For each  $\mathbf{H}^1$ , there is a  $\mu^\dagger < 1$  such that if  $\mu_\tau^2 \geq \mu^\dagger$  then for all  $t > t_2$  and  $t_1 < \tau \leq t_2$ ,  $Pr^{s,H^2}(a_t = 1) - Pr^{s,H^1}(a_t = 1)$  is decreasing in  $\zeta_\tau$ .

## A.2.2 Additional Experimental Details

Table A.1 shows the age, sex and student breakdown by arm. Chi-squared tests from multinomial a probit regression show that the randomization was successful and that none of these demographic variables are statistically different by arm:  $p = 0.20$ ,  $p = 0.41$ ,  $p = 0.29$  separately and  $p = 0.40$  jointly. The reminders and no reminders groups are also comparable: chi-square  $p = 0.40$ ,  $p = 0.23$ ,  $p = 0.10$  separately and  $p = 0.28$  jointly. And similarly for interrupted versus non-interrupted groups: chi-square  $p = 0.04$ ,  $p = 0.29$ ,  $p = 0.25$  separately and  $p = 0.15$  jointly.

**Table A.1:** *Demographics by Experimental Condition*

Condition	NB-NR	B-NR	NB-R	B-NR	Overall
Mean Age	31.3	28.1	31.2	28.4	29.6
% Male	26.0%	37.3%	23.7%	27.0%	30.5%
% Students	57.0%	58.8%	47.4%	46.0%	52.6%
Observations	54	51	38	37	180

*Notes:* Condition 1 = No interruption / No reminders; Condition 2 = Week 2 interruption / No reminders; Condition 3 = No interruption / Reminders; Condition 4 = Week 2 Interruption / Reminders.

### A.2.3 Robustness to Demographic Controls

One subject listed an out of range answer (“Beau”) in the age category, and is thus excluded from analysis. Table A.2 shows that adding demographic controls does not at all alter the results of Table 1.1. The demographic controls also do not appear to explain any additional variance in week 3 completion rates.

**Table A.2:** *Replication of Table 1.1 with Demographic Controls.*

Pr(complete survey)	(1)	(2)	(3)
Interruption	-0.420*** (0.063)	-0.419*** (0.054)	-0.421*** (0.055)
Reminders	0.034 (0.065)	0.032 (0.050)	-0.134 (0.136)
Interruption*Reminders	0.270** (0.115)	0.316*** (0.094)	0.326*** (0.095)
Week1Avg		0.635*** (0.081)	0.527*** (0.097)
Week1Avg*Reminders			0.251 (0.163)
Adj. $R^2$	0.147	0.289	0.294
Observations	1253	1253	1253

*Notes:* This table estimates a linear model of the probability of completing the daily survey in week 3 of the study, controlling for the demographic variables gathered in the registration phase of the study. The variable “Interruption” equals 1 if the daily survey was available in week 3 and equals 0 otherwise. The variable “Reminders” equals 1 if subject received daily reminders in week 3 and equals 0 otherwise. The variable “Week1avg” denotes the fraction of surveys completed in week 1. Robust standard errors clustered at subject level. All regressions include controls for: day of week, day in study, age, age<sup>2</sup>, sex, race, and whether or not subject is a student. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

#### A.2.4 Robustness to calendar date random effects

Table A.3 estimates a linear probability model with robust standard errors clustered at the subject level and calendar date level, following Cameron et al. (2011). Unfortunately, when clustering at the calendar date level, including dummy variables for day of the week and for day in study causes the variance-covariance matrix to be highly singular; thus these dummies are omitted. Instead, I just include day in study as a covariate.

**Table A.3:** Replication of Table 1.1 With Clustering at the Calendar Date Level

Pr(complete survey)	(1)	(2)	(3)
Interruption	-0.416*** (0.065)	-0.398*** (0.058)	-0.401*** (0.058)
Reminders	0.036 (0.068)	0.028 (0.056)	-0.121 (0.136)
Interruption*Reminders	0.267** (0.108)	0.305*** (0.092)	0.320*** (0.095)
Week1Avg		0.625*** (0.078)	0.527*** (0.081)
Week1Avg*Reminders			0.226 (0.152)
Adj. $R^2$	0.143	0.292	0.297
Observations	1260	1260	1260

*Notes:* This table estimates a linear probability model of completing the daily survey in week 3 of the study. The variable “Interruption” equals 1 if the daily survey was available in week 3 and equals 0 otherwise. The variable “Reminders” equals 1 if subject received daily reminders in week 3 and equals 0 otherwise. The variable “Week1avg” denotes the fraction of surveys completed in week 1. Robust standard errors clustered at the subject level and calendar date level, following Cameron et al. (2011). \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

#### A.2.5 Analysis of Post-Experimental Survey Results

Out of the 172 subjects completing the post-experimental questionnaire, 51 subjects, or 29.7% reported using a reminder technology. Subjects were coded as using a memory aid if they reported 1) using a calendar 2) using notes/diaries/daily planners/Google recurring tasks etc 3) asking others to remind them 4) leaving a tab with the study site open on the computer 5) making the study site their homepage or 6) automating their own daily reminders. Subjects were not coded as using a reminder technology if they reported a

routine such as doing the survey each morning, saying that they put it on a “mental to do list” or saying that they bookmarked the survey or starred the study overview email.

Interestingly, subjects were not less likely to use their own reminder technology if they were told that they would not be getting reminders in week 3 (Fisher’s exact test  $p = 0.40$ ). There is also no difference in reminder technology use by week 2 Interruption (Fisher’s exact test  $p = 0.17$ ).

**Table A.4:** *Effect of Using Own Reminder Technology, by Experimental Condition.*

Pr(complete survey)	(1)	(2)	(3)	(4)
Interruption	-0.432*** (0.058)	-0.438*** (0.060)	-0.458*** (0.056)	-0.463*** (0.058)
Reminders	0.051 (0.049)	-0.095 (0.136)	0.046 (0.049)	-0.110 (0.136)
Interruption*Reminders	0.338*** (0.093)	0.346*** (0.093)	0.362*** (0.100)	0.361*** (0.097)
Own_Technology	0.142*** (0.046)	0.137*** (0.045)	0.143*** (0.047)	0.137*** (0.046)
Interruption*Own_Technology	0.255*** (0.091)	0.274*** (0.092)	0.256*** (0.091)	0.276*** (0.091)
Interruption*Own_Tech*Reminders	-0.208 (0.165)	-0.209 (0.168)	-0.185 (0.163)	-0.182 (0.165)
Week1Avg	0.553*** (0.078)	0.460*** (0.090)	0.566*** (0.080)	0.467*** (0.091)
Week1Avg*Reminders		0.222 (0.162)		0.236 (0.163)
Adj. $R^2$	0.319	0.323	0.321	0.325
Observations	1204	1204	1197	1197

*Notes:* This table estimates a linear model of the probability of completing the daily survey in week 3 of the study. The variable “Interruption” equals 1 if the daily survey was available in week 3 and equals 0 otherwise. The variable “Reminders” equals 1 if subject received daily reminders in week 3 and equals 0 otherwise. The variable “Week1avg” denotes the fraction of surveys completed in week 1. The variable “Own\_Technology” equals 1 if subject reported having used a reminder technology, and equals 0 otherwise. Robust standard errors clustered at subject level. All regressions include controls for day of week and day in study. Regressions (3) and (4) also include controls for age, age<sup>2</sup>, sex, race, and whether or not subject is a student. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

## **A.3 Task Completion Experiment: Supplementary Material**

### **A.3.1 Reminders Text**

#### **Basic Email**

Dear Participant,

This is a reminder that the deadline for the risk questionnaire is [[date]], 11:59pm EST.

Thank you.

-The study team

Questions? Send an email to RiskQuestionnaires@gmail.com and we will get back to you

#### **Augmented Email**

Dear Participant,

Thank you for signing up for the risk attitudes study. This is a reminder that to receive the \$10 Amazon gift card on [[day after deadline]], you must complete the 10-20 minute risk questionnaire by [[deadline]], 11:59pm EST.

To access the questionnaire, please click on the link below and log in with the email and password with which you registered: [[Study URL]]

If you forgot your password and need help resetting it, you can reply to this email.

Thank you.

-The study team

Questions? Send an email to RiskQuestionnaires@gmail.com and we will get back to you

#### **Text Message**

This is a reminder that the deadlie for the risk questionnaire is [[date]].

### A.3.2 Demographics

Table A.5 shows demographics by experimental condition. The last row displays year in college for subjects who are students. Freshmen are coded as 1, Sophomores as 2, etc. Multinomial probit regressions do not reject the null hypothesis that student status, sex, race, and year in college are equally distributed across all 4 conditions ( $p > 0.38$  for all variables in separate regressions;  $p = 0.85$  in a joint test of significance).

**Table A.5:** *Demographics by Experimental Condition*

Condition	1	2	3	4	Overall
% Harvard students	87.1%	88.1%	89.1%	90.0%	88.6%
% Male	38.6	37.3%	23.7%	27.0%	41.7%
% White	57.0%	58.8%	47.4%	46.0%	51.6%
% Asian	29.7%	30.7%	32.7%	27.0%	30.0%
% Hispanic	6.9%	5.9%	7.9%	11.0%	7.9%
% African American	6.9%	4.0%	5.9%	2.0%	4.7%
Year (if student)	2.96	3.04	2.89	2.75	2.91
Observations	101	101	101	100	403

*Notes:* Condition 1 = Short deadline / No reminders; Condition 2 = Long deadline / No reminders; Condition 3 = Short deadline / Reminders; Condition 4 = Long deadline / Reminders.

### A.3.3 Day of week effects

**Table A.6:** *Day of Week Effects*

Pr(complete)	(1)	(2)
reminder	0.150** (0.068)	0.303*** (0.066)
Sun	-0.037 (0.138)	-0.080 (0.104)
Mon	-0.125* (0.075)	-0.135* (0.071)
Tues	0.071 (0.117)	-0.086 (0.090)
Wed	-0.000 (0.077)	-0.002 (0.134)
Thurs	-0.131 (0.121)	-0.193** (0.098)
Fri	-0.053 (0.098)	0.003 (0.092)
Adj. $R^2$	0.047	0.119
Observations	202	201
Day effects $F$ -test	$p = 0.183$	$p = 0.257$

*Notes:* This table estimates a linear probability model of completing the task, by experimental condition, to test for day of week effects. Column (1) check whether the day of week on which the deadline falls impacts completion rates for the short deadline conditions. Column2 (2) checks whether the day of week on which the deadline falls impacts completion rates for the long deadline conditions. The  $F$  tests checks the joint significance of day of week effects for each regression.

### A.3.4 Robustness

**Table A.7:** *Robustness to Day of Week and Demographics*

Pr(complete)	(1)	(2)	(3)
LongDeadline	-0.177*** (0.067)	-0.171*** (0.057)	-0.170*** (0.057)
Reminders	0.151** (0.075)	0.150** (0.076)	0.151* (0.078)
LongDeadline*Reminders	0.163* (0.089)	0.154* (0.083)	0.163* (0.091)
Adj. $R^2$	0.087	0.089	0.102
Observations	403	403	403
Demographic Controls	Yes	No	Yes
Day of Week Controls	No	Yes	Yes

*Notes:* This table estimates a linear probability model of completing the task, by experimental condition. The variable “LongDeadline” equals 1 if subjects had 3 weeks to complete the task, and equals 0 if subjects had 2 days to complete the task. The variable “Reminders” equals 1 if subjects received two days of reminders—days 1 and 2 for subjects with a 2-day deadline and days 20 and 21 for subjects with a 3-week deadline. Robust standard errors are computed by specifying both start date and undergraduate residence as the cluster groups, following the multiway clustering method suggested by Cameron et al. (2011). Demographic controls include whether or not subject is a Harvard undergraduate, sex, and race (White, Asian, Hispanic, African American, or Other). Day of week controls include dummies for which day of the week the deadline falls on. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .



### A.3.5 Reminder Types

**Table A.8:** *Different Types of Reminders Don't Have Differential Effects*

	(1)	(2)	(3)	(4)
LongDeadline	−0.178*** (0.067)	−0.178*** (0.067)	−0.178*** (0.067)	−0.178*** (0.061)
Reminders	0.158** (0.074)	0.163** (0.078)	0.165 (0.106)	0.164* (0.088)
LongDeadline*Reminders	0.156* (0.081)	0.146 (0.091)	0.154* (0.082)	0.157* (0.086)
Info	−0.031 (0.052)	−0.047 (0.112)		
Info*LongDeadline		0.033 (0.176)		
SMSmessage			−0.025 (0.073)	−0.022 (0.069)
SMS*LongDeadline				−0.005 (0.116)
Adj. $R^2$	0.074	0.074	0.074	0.060
Observations	403	403	403	403

*Notes:* This table estimates a linear probability model of completing the task, by experimental condition. The variable “LongDeadline” equals 1 if subjects had 3 weeks to complete the task, and equals 0 if subjects had 2 days to complete the task. The variable “Reminders” equals 1 if subjects received two days of reminders—days 1 and 2 for subjects with a 2-day deadline and days 20 and 21 for subjects with a 3-week deadline. The variable “Info” equals 1 if the email message contained information in addition to the deadline. The variable SMSMessage equals 1 if subject request to receive SMS message reminders in addition to the emails. Robust standard errors are computed by specifying both start date and undergraduate residence as the cluster groups, following the multiway clustering method suggested by Cameron et al. (2011).

## A.4 Proofs of Mathematical Results

### A.4.1 Proofs for repeated action results

As before, I will let  $s_t^s$  and  $s_t^n$  denote the sophisticated and naive DMs’ threshold strategies, respectively: the DM chooses  $x_t = a$  if and only if  $\zeta_t \geq s_t$ . I will let  $\mathbf{s}^s = (s_1^s, \dots, s_T^s)$  denote the sophisticated DM’s vector of threshold strategies, and I will let  $\mathbf{s}^n$  denote the naive DM’s vector of threshold strategies.

Since past of actions does not affect payoffs in this setting, I will use  $V_t^s(\alpha_t)$  and  $V_t^n(\alpha_t)$

to denote the sophisticated and naive DMs' expected period  $t$  utility conditional on being attentive or not that period.

I will let  $Pr(\alpha_t = 1)$  denote the probability of being attentive in period  $t$ , from the period 0 perspective.

### Preliminary results

I begin by showing that for both sophisticated and naive DMs,  $Pr(x_t = a)$  is differentiable in  $b_{t'}$  for any  $t' \geq \tau$ .

**Lemma 3.**  $Pr^n(x_t = a)$  is differentiable in  $b_{t'}$  for any  $t' \geq \tau$ .

*Proof.* Note that this is clearly true for  $t < t'$ , since a naive DM does not adjust his current behavior in anticipation of future benefits.

Now suppose that  $t = t'$ . Because  $Pr^n(x_t = a) = Pr(b_t + \xi_t \geq 0)Pr^n(\alpha_t = 1)$ , it follows that  $Pr^n(x_t = a) = Pr(b_t + \xi_t \geq 0)Pr^n(\alpha_t = 1)$ . But  $Pr(b_t + \xi_t \geq 0) = 1 - F(-b_t)$  is of course differentiable in  $b_t$  by assumption.

Next suppose that  $t > t'$ . For  $t > t'$ ,

$$\begin{aligned} Pr^n(\alpha_{t+1} = 1) &= \gamma_{t+1}(0, d) + Pr^n(\alpha_t = 1)(\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d)) \\ &+ Pr^n(\alpha_t = 1)Pr(b_t + \xi_t \geq 0)(\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)). \end{aligned}$$

Thus  $Pr^n(\alpha_{t+1} = 1)$  is differentiable in  $b_{t'}$  if  $Pr^n(\alpha_t)$  and  $Pr(b_t + \xi_t \geq 0)$  are differentiable in  $b_{t'}$ . A simple induction thus shows that  $Pr^n(\alpha_t = 1)$  is differentiable in  $b_{t'}$  for all  $t > t'$ . And because  $Pr^n(x_t = a) = Pr(b_t + \xi_t \geq 0)Pr^n(\alpha_t = 1)$ , it then follows that  $Pr^n(x_t = a)$  is differentiable in  $b_{t'}$  for all  $t > t'$ .  $\square$

**Lemma 4.**  $Pr^s(x_t = a)$  is differentiable in  $b_{t'}$  for any  $t' \geq \tau$ .

*Proof.* First, suppose that  $t = t'$ . Then  $V_t^s(0)$ ,  $V_{t+1}^s(1)$ , and  $V_{t+1}^s(0)$  are all differentiable in  $b_{t'}$

because they are not functions of  $b_{t'}$ . By definition,

$$s_t^s = -b_t - [\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)] \Delta V_{t+1}^s \quad (\text{A.3})$$

$$V_t^s(0) = V_{t+1}^s(0) + \gamma_{t+1}(0, d) \Delta V_{t+1}^s \quad (\text{A.4})$$

$$V_t^s(1) = V_{t+1}^s(0) + \gamma_{t+1}(1, d) \Delta V_{t+1}^s + \int_{\xi_t \geq s_t^s} [(b_t + \xi_t) + (\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)) \Delta V_{t+1}^s] \lambda d\xi \quad (\text{A.5})$$

where  $\Delta V_{t+1}^s = V_{t+1}^s(1) - V_{t+1}^s(0)$ . Now suppose that  $t = t'$ . Then  $V_t^s(0)$ ,  $V_{t+1}^s(1)$ , and  $V_{t+1}^s(0)$  are all differentiable in  $b_{t'}$  because they are not functions of  $b_{t'}$ , and thus equations (A.3)-(A.5) show that  $s_t^s$  and  $V_t^s(1)$  are differentiable in  $b_t$ . By recursive reasoning, equations (A.3)-(A.5) more generally show that  $s_t^s$  and  $V_t^s(\alpha_t)$  are differentiable in  $b_{t'}$  for  $t \leq t'$ . Moreover,  $s_t^s$  is differentiable in  $b_{t'}$  for  $t > t'$  for the simple reason that  $s_t^s$  is not a function of  $b_{t'}$  for  $t > t'$ .

Next, note that

$$\begin{aligned} Pr^s(\alpha_{t+1} = 1) &= \gamma_{t+1}(0, d) + Pr^s(\alpha_t = 1)(\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d)) \\ &+ Pr^s(\alpha_t = 1)Pr(\xi_t \geq s_t)(\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)). \end{aligned} \quad (\text{A.6})$$

from which it follows that  $Pr^s(\alpha_{t+1} = 1)$  is differentiable in  $b_{t'}$  if  $Pr^s(\alpha_t = 1)$  is differentiable in  $b_{t'}$ . Similarly, (A.6) also shows that  $Pr^s(\alpha_t = 1)$  is differentiable in  $b_{t'}$  if  $Pr^s(\alpha_{t+1} = 1)$  is differentiable in  $b_{t'}$ . But because  $Pr(\alpha_t)$  is differentiable in  $b_{t'}$  for  $t = t'$ , recursive reasoning shows that  $Pr(\alpha_t)$  is differentiable in  $b_{t'}$  for all  $t$ .

Finally, because  $Pr^n(x_t = a) = Pr(\xi_t \geq s_t^s)Pr^n(\alpha_t = 1)$ , it then follows that  $Pr^n(x_t = a)$  is differentiable in  $b_{t'}$  for all  $t$ .  $\square$

## Proofs of Propositions

*Proof of Proposition 1, parts 1 and 2.* A naive DM chooses  $x_{t'} = a$  if and only if  $b_{t'} + \xi_{t'} \geq 0$ . So an increase in  $b_{t'}$  clearly increases the likelihood of choosing  $x_{t'} = a$ . Because of the behavioral rehearsal property, this increases the probability of being attentive in period  $t + 1$ , and subsequently choosing  $x_{t'+1} = a$ . A simple induction shows that for all periods  $t > t'$  the DM will thus be more likely to be attentive, and therefore choose  $x_t = a$ .

I now turn to sophisticated DMs. I first consider how  $s^s$  changes as  $b_{t'}$  changes. Conditional on being attentive in some period  $t > t'$ , a change in  $b_{t'}$  does not affect the DM's payoff in any period  $\tau \geq t$ , or the likelihood of the DM being attentive in any period  $\tau > t$ . Thus  $s_t^s$  remains fixed for  $t > t'$ .

Similarly,  $V_t^s(\alpha_t)$  remains fixed for any  $t > t'$ . The DM's optimal strategy in any period  $t$  takes into account period  $t$  flow utility, as well as how it affects future expected utility through changes in future expected attentiveness,

$$s_t^s = -b_t - [\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)] \Delta V_{t+1}^s \quad (\text{A.7})$$

where  $\Delta V_{t+1}^s = V_{t+1}^s(1) - V_{t+1}^s(0)$ . But since  $V_t^s$  is constant in  $b_{t'}$  for all  $t > t'$ , it follows that  $s_{t'}^s$  is decreasing in  $b_{t'}$ .

Next, for  $s_t^s$  defined as in (A.7),

$$\begin{aligned} V_t^s(1) - V_t^s(0) &= V_{t+1}^s(0) + \gamma_{t+1}(1, d) \Delta V_{t+1}^s + \int_{\xi_t \geq s_t^s} [(b_t + \xi_t) + (\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)) \Delta V_{t+1}^s] dF \\ &\quad - [V_{t+1}^s(0) + \gamma_{t+1}(0, d) \Delta V_{t+1}^s] \\ &= \int_{\xi_t \geq s_t^s} [(b_t + \xi_t) + (\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)) \Delta V_{t+1}^s] dF \\ &\quad + (\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d)) \Delta V_{t+1}^s. \end{aligned}$$

Differentiating  $V_t^s(1) - V_t^s(0)$  with respect to  $b_t$ , shows that

$$\begin{aligned} \frac{\partial}{\partial b_t} (V_t^s(1) - V_t^s(0)) &= \int_{\xi_t \geq s_t^s} (1) dF + [(b_t + s_t^s) + (\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)) \Delta V_{t+1}^s] f(s_t^s) \\ &= 1 - F(s_t^s) > 0 \end{aligned} \quad (\text{A.8})$$

where (A.7) is used to obtain the second line. This shows that  $\Delta V_{t'}^s$  is increasing in  $b_{t'}$ .

Next, differentiating  $V_t^s(1) - V_t^s(0)$  with respect to  $\Delta V_{t+1}^s$ , shows that

$$\begin{aligned} \frac{d}{d\Delta V_{t+1}^s} (V_t^s(1) - V_t^s(0)) &= \int_{\xi_t \geq s_t^s} (\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)) dF + [(b_t + s_t^s) + (\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)) \Delta V_{t+1}^s] f(s_t^s) \\ &= 1 - F(\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)) > 0 \end{aligned} \quad (\text{A.9})$$

Thus, since  $\Delta V_{t'}^s$  is increasing in  $b_{t'}$ , a straightforward induction shows that  $\Delta V_t^s$  is increasing in  $b_{t'}$  for  $t < t'$ . Equation (A.7) thus shows that  $s_t^s$  is decreasing in  $b_{t'}$  for all  $t \leq t'$ .

Equation (A.6) shows that  $Pr^s(\alpha_{t+1} = 1)$  is increasing in  $Pr^s(\alpha_t = 1)$  and  $Pr(\xi_t \geq s_t)$ . And because  $Pr^s(\alpha_1 = 1)$  is constant in  $b_{t'}$ , recursively applying equation (A.6) shows that  $Pr^s(\alpha_{t'} = 1)$  is increasing in  $b_t$  for all  $t'$ .  $\square$

*Proof of Proposition 1, part 3 (naive DM).* I first consider naive DMs. Note that  $Pr^{n,Hi}(x_t = a) = Pr(b_t + \xi_t \geq 0)Pr^{n,Hi}(\alpha_t = 1)$ . Thus

$$\frac{dPr^{n,Hi}(x_t = a)}{db_{t'}} = \frac{dPr(b_t + \xi_t \geq 0)}{db_{t'}} Pr^{n,Hi}(\alpha_t = 1).$$

So the result follows for  $t' = t$  if  $Pr^n(\alpha_t = 1)$  is increasing  $b_{t'}$ . This is true, as shown in the proof of Proposition 1, part 1.

Now

$$\begin{aligned} Pr^n(\alpha_{t+1} = 1) &= \gamma_{t+1}(0, d) + Pr^n(\alpha_t = 1)(\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d)) \\ &+ Pr^n(\alpha_t = 1)Pr^n(b_t + \xi_t \geq 0)(\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)), \end{aligned} \quad (\text{A.10})$$

which shows that

$$\frac{\partial Pr^n(\alpha_{t+1} = 1)}{\partial b_t} = Pr^n(\alpha_t = 1) \frac{\partial Pr^n(b_t + \xi_t \geq 0)}{\partial b_t}$$

is increasing in  $Pr^n(\alpha_t = 1)$ . For  $t' > t$ , similar computations show that

$$\begin{aligned} \frac{\partial Pr^n(\alpha_{t'+1} = 1)}{\partial b_t} &= \frac{\partial Pr^n(\alpha_{t'} = 1)}{\partial b_t} (\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d)) \\ &+ \frac{\partial Pr^n(\alpha_{t'} = 1)}{\partial b_t} Pr^n(b_{t'} + \xi_{t'} \geq 0) (\gamma_{t'+1}(1, a) - \gamma_{t'+1}(1, d)) \end{aligned}$$

A simple proof by induction then shows that  $\frac{dPr^n(\alpha_{t'}=1)}{db_t}$  is increasing in  $Pr^n(\alpha_t = 1)$  for all  $t' > t$ . But since  $Pr^n(\alpha_t = 1)$  is increasing in  $b_{t'}$ , the statement of the proposition follows.  $\square$

*Proof of Proposition 1, part 3 (sophisticated DM).* Equation (A.8) shows that

$$\frac{\partial}{\partial b_{t'}} (V_{t'}^s(1) - V_{t'}^s(0)) = 1 - F(s_{t'}^s) \quad (\text{A.11})$$

from which it follows that

$$\frac{\partial^2}{\partial b_{t''} \partial b_t} (V_{t''}^s(1) - V_{t''}^s(0)) = -f(s_{t''}^s) \frac{\partial}{\partial b_t} s_{t''}^s > 0 \quad (\text{A.12})$$

because  $\frac{\partial}{\partial b_t} s_{t''}^s < 0$ , as established in part 2. Equation (A.9) combined with a simple induction now shows that

$$\frac{\partial^2}{\partial b_{t''} \partial b_t} (V_\tau^s(1) - V_\tau^s(0)) > 0 \quad (\text{A.13})$$

for all  $\tau \leq t''$ . Equation (A.7) now shows that

$$\frac{\partial^2}{\partial b_{t''} \partial b_t} s_\tau^s < 0 \quad (\text{A.14})$$

for all  $\tau < t''$  and that

$$\frac{\partial^2}{\partial b_{t''} \partial b_t} s_{t''}^s = 0 \quad (\text{A.15})$$

for  $\tau \geq t''$ .

Now equation (A.6), implies that for any  $t' \geq \tau$ ,

$$\frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\alpha_{t'+1}) = \frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\alpha_{t'}) (\gamma_{t'+1}(1, d) - \gamma_{t'+1}(0, d)) \quad (\text{A.16})$$

$$+ \left[ \left( \frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\xi_{t'} \geq s_{t'}) \right) Pr^s(\alpha_{t'}) \right] \quad (\text{A.17})$$

$$+ Pr^s(\xi_{t'} \geq s_{t'}) \left( \frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\alpha_{t'}) \right) \quad (\text{A.18})$$

$$+ \left( \frac{\partial}{\partial b_{t''}} Pr^s(\xi_{t'} \geq s_{t'}) \right) \left( \frac{\partial}{\partial b_t} Pr^s(\alpha_{t'}) \right) \quad (\text{A.19})$$

$$+ \left( \frac{\partial}{\partial b_t} Pr^s(\xi_{t'} \geq s_{t'}) \right) \left( \frac{\partial}{\partial b_{t''}} Pr^s(\alpha_{t'}) \right) \quad (\text{A.20})$$

from which it follows that  $\frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\alpha_{t'+1}) > 0$  if  $\frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\alpha_{t'}) \geq 0$ . But since  $\frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\alpha_1) \geq 0$ , recursively using the computations above, as well as equations (A.14) and (A.15) shows that  $\frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\alpha_{t'}) > 0$  for all  $t' \leq t$ . Reasoning analogous to part 1 now shows that  $\frac{\partial^2}{\partial b_{t''} \partial b_t} Pr^s(\alpha_{t'}) > 0$  for all  $t' > t$  as well.  $\square$

*Proof of Proposition 2, parts 1,2.* Part 2 is obvious because naive DMs just choose  $x_t = a$  if  $b_t + \xi_t \geq 0$ . Part 1 follows because the DM will be more attentive in period  $t_2$ , which, as in

the proof of Proposition 1 (part 1), will make him more likely to be attentive in all future periods.  $\square$

*Proof of Proposition 2, part 3.* That the result holds under condition (a) is obvious, and follows from the same logic as for the naive case.

Consider now condition (b). Assumptions A1 and A2 imply that there exists an  $\epsilon$  such that  $\int g(\alpha, x, \sigma) dH_{t'}^1 < (1 - \epsilon)g(\alpha, x, 1 - \epsilon)$ . But now if  $\mu_{t'}^2 > 1 - \epsilon^2$ , then under  $H_{t'}^2$ ,  $\sigma > 1 - \epsilon$  with probability at least  $1 - \epsilon$ . Thus for a high enough  $\mu_{t'}^2$ ,  $Pr^{s, H^2}(\alpha_{t'} = 1) > Pr^{s, H^1}(\alpha_{t'} = 1)$ . From this, it follows that the probability of attentiveness under  $H^2$  is higher in all period  $t > t'$  as well. Moreover, as in the proof of Proposition 1, the sophisticated DM's strategies don't change for  $t \geq t'$ , from which the results follow.  $\square$

*Proof of Proposition 2, part 4.* Let  $\gamma_t^i(\alpha_{t-1}, x_{t-1})$  denote the probability of being attentive in period  $t$  under cue distribution  $\mathbf{H}_t^i$ . As shown in the proof of part 2 of Proposition 1,

$$V_t^{s,i}(1) - V_t^{s,i}(0) = \int_{\xi_t \geq s_t^s} \left[ (b_t + \xi_t) + \left( \gamma_{t+1}^i(1, a) - \gamma_{t+1}^i(1, d) \right) \Delta V_{t+1}^s \right] dF + \left( \gamma_{t+1}^i(1, d) - \gamma_{t+1}^i(0, d) \right) \Delta V_{t+1}^s$$

Consider now  $t = t' - 1$ . By assumption A4,  $\gamma_{t+1}^2(1, a) - \gamma_{t+1}^2(1, d) < \gamma_{t+1}^1(1, a) - \gamma_{t+1}^1(1, d)$  and  $\gamma_{t+1}^2(1, d) - \gamma_{t+1}^2(0, d) < \gamma_{t+1}^1(1, d) - \gamma_{t+1}^1(0, d)$ . It thus follows that  $V_{t+1}^{s,2}(1) - V_{t+1}^{s,2}(0) < V_{t+1}^{s,1}(1) - V_{t+1}^{s,1}(0)$ . Now the proof of part 2 of Proposition 1 shows that  $V_t^s(1) - V_t^s(0)$  is increasing in  $\Delta V_{t+1}^s$ . From this it follows that  $V^{s,2}(1) - V^{s,2}(0) < V^{s,1}(1) - V^{s,1}(0)$  for all  $t = t_1 + 1, t_1 + 1, \dots, t_2$ . As shown in equation (A.7), this implies that  $s_t^s$  is higher under  $\mathbf{H}^2$ ; which means that the DM will be less likely to do the task under  $\mathbf{H}^2$ .  $\square$

*Proof of Proposition 2, part 5.* By definition,

$$Pr^{n, H^2}(\alpha_{t_2} = 1) - Pr^{n, H^1}(\alpha_{t_2} = 1) \tag{A.21}$$

$$= \left( \gamma_{t_2}^2(0, d) - \gamma_{t_2}^1(0, d) \right) + Pr^n(\alpha_{t_2-1} = 1) \left[ \left( \gamma_{t_2}^2(1, d) - \gamma_{t_2}^1(1, d) \right) - \left( \gamma_{t_2}^2(0, d) - \gamma_{t_2}^1(0, d) \right) \right] \tag{A.22}$$

$$+ Pr^n(\alpha_{t_2-1} = 1) Pr(b_{t_2-1} + \xi_{t_2-1} \geq 0) \left[ \left( \gamma_{t_2}^2(1, a) - \gamma_{t_2}^1(1, a) \right) - \left( \gamma_{t_2}^2(1, d) - \gamma_{t_2}^1(1, d) \right) \right]. \tag{A.23}$$

But assumption A4' implies that that  $(\gamma_{t_2}^2(1, d) - \gamma_{t_2}^1(1, d)) - (\gamma_{t_2}^2(0, d) - \gamma_{t_2}^1(0, d)) < 0$  and that  $(\gamma_{t_2}^2(1, a) - \gamma_{t_2}^1(1, a)) - (\gamma_{t_2}^2(1, d) - \gamma_{t_2}^1(1, d)) < 0$ . It thus follows that  $Pr^{n, H^2}(\alpha_{t_2} = 1) - Pr^{n, H^1}(\alpha_{t_2} = 1)$  is decreasing in  $Pr^n(\alpha_{t_2-1} = 1)$ . But since  $Pr^n(\alpha_{t_2-1} = 1)$  is increasing in  $H_{t'}$  by part 1, it thus follows that  $Pr^{n, H^2}(\alpha_{t_2} = 1) - Pr^{n, H^1}(\alpha_{t_2} = 1)$  is decreasing in  $H_{t'}$ .

Finally, note that for  $t > t_2$ ,

$$Pr^{n, H^2}(\alpha_t = 1) - Pr^{n, H^1}(\alpha_t = 1) \tag{A.24}$$

$$= \left( Pr^{n, H^2}(\alpha_t = 1) - Pr^{n, H^1}(\alpha_t = 1) \right) (\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d)) \tag{A.25}$$

$$= \left( Pr^{n, H^2}(\alpha_t = 1) - Pr^{n, H^1}(\alpha_t = 1) \right) Pr(b_t + \xi_t \geq 0) (\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)) \tag{A.26}$$

is increasing in  $Pr^{n, H^2}(\alpha_t = 1) - Pr^{n, H^1}(\alpha_t = 1)$ . A simple proof by induction thus shows that  $Pr^{n, H^2}(\alpha_t = 1) - Pr^{n, H^1}(\alpha_t = 1)$  is decreasing in  $H_{t'}$  for all  $t \geq t_2$ .  $\square$

*Proof of Proposition 3, part 1.* Note that  $Pr^{n, H^i}(x_t = a) = Pr(b_t + \xi_t \geq 0) Pr^{n, H^i}(\alpha_t = 1)$ . Thus

$$\frac{\partial Pr^{n, H^i}(x_t = a)}{\partial b_{t_2}} = \frac{\partial Pr(b_t + \xi_t \geq 0)}{\partial b_{t_2}} Pr^{n, H^i}(\alpha_t = 1).$$

So the result follows for  $t = t_2$  if  $Pr^{n, H^2}(\alpha_{t_2} = 1) > Pr^{n, H^1}(\alpha_{t_2} = 1)$ . This is true, as shown in the proof of Proposition 2, part 1.

Now for the naive DM,

$$\begin{aligned} Pr(\alpha_{t+1} = 1) &= \gamma_{t+1}(0, d) + Pr(\alpha_t = 1)(\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d)) \\ &+ Pr(\alpha_t = 1)Pr(b_t + \xi_t \geq 0)(\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d)), \end{aligned} \tag{A.27}$$

which shows that  $\frac{\partial Pr(x_{t+1}=a)}{\partial b_{t_2}}$  is increasing in  $Pr(\alpha_t = 1)$  when  $t = t_2$ . A simple proof by induction then shows that  $\frac{\partial Pr(x_{t+1}=a)}{\partial b_{t_2}}$  is increasing in  $Pr(\alpha_t = 1)$  for  $t \geq t_2$ . Since,  $Pr^{n, H^2}(\alpha_{t_2} = 1) > Pr^{n, H^1}(\alpha_{t_2} = 1)$ , the statement of the proposition follows.  $\square$

*Proof of Proposition 3, part 2.* When looking at period 1 behavior, there is no scope for cues to crowd out behavioral rehearsal as in Proposition 2, part 4. Moreover, period  $t_1$  cues do not affect the sophisticated DM's strategy in periods  $t \geq t_1$ . Thus the rest of the argument follows analogously to the argument for part 1.  $\square$



*Proof of Proposition 4, part 1.* As shown in equation (A.27), the impact of period  $t$  rehearsal on period  $t + 1$  attentiveness is increasing in  $(\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d))$  and  $(\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d))$ . Assumption A4 guarantees that  $g(1, d, \sigma) - g(0, d, \sigma)$  are decreasing in  $\sigma$ , from which it follows that

$$(\gamma_{t+1}(1, d) - \gamma_{t+1}(0, d)) = \int (g(1, d, \sigma) - g(0, d, \sigma)) dH_{t+1}$$

is decreasing in  $H_{t+1}$  (in the FOSD order). Similarly,  $(\gamma_{t+1}(1, a) - \gamma_{t+1}(1, d))$  are decreasing  $H_{t+1}$ . Thus the impact of rehearsal in periods  $t < t_2$  on the likelihood of being attentive in period  $t = t_2$  is smaller under  $\mathbf{H}^2$ .  $\square$

*Proof of Proposition 4, part 2.* Assumptions A1 and A2 guarantee that  $\gamma_{t_2}(0, d) \rightarrow 1$  as  $\mu_{t_2}^2 \rightarrow 1$ . This implies that the impact of period  $t < t_2$  rehearsal on period  $t_2$  attentiveness can be made arbitrarily small when  $\mu_{t_2}^2$  is sufficiently high. As a consequence, the impact of period  $t < t_2$  rehearsal on period  $t' > t_2$  attentiveness can be made arbitrarily small when  $\mu_{t_2}^2$  is sufficiently high.  $\square$

#### A.4.2 Proofs for one time actions results

Throughout, I will use the following additional notation. I will use  $s_t^{\text{pa}, T}$  to denote the sophisticated DM's period  $t$  threshold rule, given a deadline  $T$ . That is, the sophisticated DM chooses  $x_t = a$  if and only if  $\zeta_t \geq s_t^{\text{pa}, T}$ . I will define  $s_t^{s, T}$  and  $s_t^{n, T}$  similarly. When considering results under two different deadlines  $T_1$ , and  $T_2$ , I will use the superscript  $T_i$  to index the respective strategies  $s_t^{\text{pa}, T_i}$ ,  $s_t^{s, T_i}$ ,  $s_t^{n, T_i}$ . I will let  $\mathbf{d}_t$  denote the  $1 \times t$  vector  $(d, \dots, d)$ . Then  $V_t^{\text{pa}}(\mathbf{d}_{t-1})$  will denote the perfectly attentive DM's period  $t$  utility, conditional on not having yet completed the task.

#### Preliminary Results

I begin with several lemmas that will be used in the proofs of the propositions.

**Lemma 5.** *Suppose that  $b_t \equiv b$  for all  $t$ . Then  $V_0^{\text{pa}, T} \rightarrow b + \bar{\zeta}$  as  $T \rightarrow \infty$ .*

*Proof.* Suppose that the DM follows the following strategy: For some  $\epsilon > 0$ , choose  $x_t = a$  if and only if  $b + \zeta_t > b + \bar{\zeta} - \epsilon/2$ . Then conditional on completing the task, the DM's utility is at least  $b + \bar{\zeta} - \epsilon/2$ . And his probability of completing the task is at least  $p_{\epsilon,T} = 1 - (1 - F(\bar{\zeta} - \epsilon/2))^T$ . Clearly,  $p_{\epsilon,T} \rightarrow 1$  as  $T \rightarrow \infty$ . Thus for any  $\epsilon > 0$ ,

$$\lim_{T \rightarrow \infty} V_0^{\text{pa},T} > b + \bar{\zeta} - \epsilon. \quad (\text{A.28})$$

But since equation (A.28) holds for any  $\epsilon > 0$ , it therefore follows that  $V_0^{\text{pa},T} \rightarrow b + \bar{\zeta}$  as  $T \rightarrow \infty$ .  $\square$

**Lemma 6.** *Suppose that  $b_t \equiv b$  for all  $t$ . Then  $\Pr(Q^{\text{pa},T} \leq T) \rightarrow 1$  as  $T \rightarrow \infty$ .*

*Proof.* Suppose, by way of contradiction, that  $1 - \Pr(Q^{\text{pa},T} \leq T) > \eta$  for some  $\eta > 0$ . This would then imply that  $V_0^{\text{pa},T} < (b + \bar{\zeta})(1 - \eta)$  for all  $T$ . This is in direct contradiction to Lemma 5.  $\square$

**Lemma 7.** *Suppose that  $b_t \equiv b$  for all  $t$ . Then for any fixed  $t$ ,  $V_t^{\text{pa},T}(\mathbf{d}_{t-1}) \rightarrow b + \bar{\zeta}$  as  $T \rightarrow \infty$ .*

*Proof.* Follows identically to the proof of Lemma 5.  $\square$

**Lemma 8.** *Fix some  $t \geq 1$ . Then  $\lim_{T \rightarrow \infty} s_t^{\text{pa},T} = \bar{\zeta}$ .*

*Proof.* By definition,  $b + s^{\text{pa},T} = V_{t+1}^{\text{pa},T}$ . By Lemma 7,  $V_{t+1}^{\text{pa},T}(\mathbf{d}_{t-1}) \rightarrow b + \bar{\zeta}$  as  $T \rightarrow \infty$ . Thus  $s^{\text{pa},T} \rightarrow \bar{\zeta}$  as  $T \rightarrow \infty$ .  $\square$

**Lemma 9.** *For some  $t^\dagger \leq T$ , set  $b_t = b'_t + \eta$  for  $t \leq t^\dagger$ , and  $b_t = b'_t$  for  $t > t^\dagger$ . Then  $\Pr(Q^{\text{pa}} \leq T)$  is increasing in  $\eta$ .*

*Proof.* I will let  $V_t^{\text{pa}}(\mathbf{d}_{t-1}; \eta)$  denote the expected period  $t$  utility as a function of  $\eta$ . For  $t > t^\dagger$ ,  $V_t^{\text{pa}}(\mathbf{d}_{t-1}; \eta)$  does not depend on  $\eta$ . For  $t \leq t^\dagger$ , I will now show that  $V_t^{\text{pa}}(\mathbf{d}_{t-1}; \eta) - V_t^{\text{pa}}(\mathbf{d}_{t-1}; 0) \leq \eta$ . Note that the threshold rule is given by  $s_t^s(\eta) = \max\left(V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1}) - b'_t - \eta, \bar{\zeta}\right)$ .

Setting  $V_{T+1}(\mathbf{d}_T) = 0$ , we now have:

$$\begin{aligned}
V_t^{\text{pa}}(\mathbf{d}_{t-1}; \eta) &= \int_{s_t^s(\eta)}^{\bar{x}i} (b'_t + \eta - \zeta_t) dF + F(s_t^s(\eta)) V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1}; \eta) \\
&= \int_{s_t^s(0)}^{s_t^s(\eta)} (b'_t + \eta - \zeta_t) dF + \int_{s_t^s(0)}^{\bar{x}i} (b'_t + \eta - \zeta_t) dF + F(s_t^s(\eta)) V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1}; \eta) \\
&= \int_{s_t^s(0)}^{s_t^s(\eta)} [(b'_t + \eta - \zeta_t) - V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1})] dF + \int_{s_t^s(0)}^{\bar{x}i} (b'_t + \eta - \zeta_t) dF + F(s_t^s(0)) V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1}; \eta) \\
&= (1 - F(s_t^s(\eta))) \eta + \int_{s_t^s(0)}^{s_t^s(\eta)} [(b'_t - \zeta_t) - V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1})] dF + \int_{s_t^s(0)}^{\bar{x}i} (b'_t - \zeta_t) dF + F(s_t^s(0)) V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1}; \eta) \\
&\leq (1 - F(s_t^s(\eta))) \eta + \int_{s_t^s(0)}^{\bar{x}i} (b'_t - \zeta_t) dF + F(s_t^s(0)) V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1}; \eta) \\
&= (1 - F(s_t^s(\eta))) \eta + F(s_t^s(0)) [V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1}; \eta) - V_{t+1}^{\text{pa}}(\mathbf{d}_{t-1}; 0)] + V_t^{\text{pa}}(\mathbf{d}_{t-1}; 0).
\end{aligned}$$

Now clearly,  $V_{T+1}^{\text{pa}}(\mathbf{d}_T; \eta) - V_{T+1}^{\text{pa}}(\mathbf{d}_T; 0) < \eta$ , and thus a simple induction shows that  $V_t^{\text{pa}}(\mathbf{d}_{t-1}; \eta) - V_t^{\text{pa}}(\mathbf{d}_{t-1}; 0) \leq \eta$ .  $\square$

### Proof of Propositions

*Proof of Proposition 5. Part 1.* Obvious. Conditional on being attentive, the naive DM follows the same strategy as the perfectly attentive DM. However, the naive DM is less likely to do the task each period because he is inattentive with some positive probability.

*Part 2.* Clearly,  $\Pr(Q^s \leq T) < \Pr(Q^{\text{pa}} \leq T)$  when  $b + \underline{\zeta} \geq 0$ : by assumption, the perfectly attentive DM will complete the task with probability 1, whereas the inattentive DM will complete the task with probability less than one. Now by assumption, the inattentive DM is attentive with probability no greater than  $\bar{z} < 1$  in all periods, and thus  $V_2^s \leq \bar{z}(b + \bar{\zeta})$  for any value of  $T$ . This means that  $s_1^s$  is bounded away from  $\bar{\zeta}$  for all  $T$  or, equivalently, that  $\Pr^s(x_1 = a)$  is bounded away from 0 for all  $T$ . On the other hand, Lemma 8 implies that for any  $t^\dagger$ ,  $\Pr(Q^{\text{pa}} \leq t^\dagger) \rightarrow 0$  as  $T \rightarrow \infty$ .

*Part 3.* Obvious. Sufficiently low probability of being attentive can make completion probabilities arbitrarily low.  $\square$

*Proof of Proposition 6. Part 1.* Obvious

*Part 2.* An even stronger result is true. Under condition  $I(T_1, T_2)$ , the probability of

completing the task between periods  $t = \Delta + 1$  and  $t = T_2$ , conditional on not having completed the task by period  $\Delta + 1$ , is higher than the probability of completing the task when given the short deadline. Formally,  $Pr(Q^{\text{pa},T_1} \leq T_1) < Pr(Q^{\text{pa},T_2} \leq T_2 | Q^{\text{pa},T_2} > \Delta)$ .

The short deadline game can be transformed into the  $T_1$  period subgame of the long deadline through the following series of operations:

- Increase the payoff in all periods by  $b_{T_2} - b_{T_1}$ .
- Increase the payoff all but the last period by  $(b_{T_2-1} - b_{T_1-1}) - (b_{T_2} - b_{T_1})$
- ...
- Increase the payoff in the first period by  $(b_{\Delta+1} - b_1) - (b_{\Delta+2} - b_2)$ .

Lemma 9 implies that each operation increase the likelihood of completing the task, which establishes the result.

*Part 3.* Note that  $V_t^{s,T_i} \leq \bar{z}(b + \bar{\zeta})$  for all  $t$ . By definition,  $s_t^{s,T_i} = \max(V_{t+1}^{s,T_i} - b, \underline{\zeta})$ , from which it follows that  $s_t^{s,T_i} = \max(\bar{z}\bar{\zeta} - (1 - \bar{z})b, \underline{\zeta} - b)$ . However,  $\bar{z}\bar{\zeta} - (1 - \bar{z})b \rightarrow -\infty$  as  $b \rightarrow \infty$ , from which it follows that  $s_t^{s,T_i} \rightarrow \underline{\zeta}$  as  $b \rightarrow \infty$ .

Thus for  $t \leq T_1$ ,  $s_t^{s,T_1} / s_t^{s,T_2} \rightarrow 1$  as  $b \rightarrow \infty$ . Moreover, the assumption that  $\gamma_t(1, d)$  is bounded away from 1 implies that  $\lim_{b \rightarrow \infty} Pr(Q^{s,T_i} \leq T_i) < 1$  each  $i = 1, 2$ . Thus it follows that

$$\begin{aligned} \lim_{b \rightarrow \infty} 1 - Pr(Q^{s,T_2} \leq T_2) &= \lim_{b \rightarrow \infty} Pr(T_1 + 1 \leq Q^{s,T_2} \leq T_2) \left(1 - Pr(Q^{s,T_2} \leq T_1)\right) \\ &< \lim_{b \rightarrow \infty} \left(1 - Pr(Q^{s,T_2} \leq T_1)\right), \end{aligned}$$

which completes the proof. □

*Proof of Proposition 7.* Clearly,  $V_t^{\text{pa},T_2} > V_t^{\text{pa},T_1}$  for all  $t \leq T_1$ . Therefore,  $s_t^{\text{n},T_1} < s_t^{\text{n},T_2}$  for all  $t \leq T_1$ . This implies that  $Pr(2 \leq Q^{\text{n},T_2} \leq T_1) < Pr(2 \leq Q^{\text{n},T_1} \leq T_1)$ . Noting that  $\gamma_1(1, d)$  is

the probability of being attentive in period 1 under both deadlines,

$$\begin{aligned}
Pr(Q^{n,T_1} \leq T_1) - Pr(Q^{n,T_2} \leq T_1) &= \left[1 - \gamma_1(1,d)(1 - F(s_1^{s,T_2}))\right] \left[1 - Pr(2 \leq Q^{n,T_2} \leq T_1 | Q^{n,T_2} > 1)\right] \\
&- \left[1 - \gamma_1(1,d)(1 - F(s_1^{s,T_1}))\right] \left[1 - Pr(2 \leq Q^{n,T_1} \leq T_1 | Q^{n,T_1} > 1)\right] \\
&> \left[1 - \gamma_1(1,d)(1 - F(s_1^{s,T_2}))\right] \left[1 - Pr(2 \leq Q^{n,T_2} \leq T_1 | Q^{n,T_2} > 1)\right] \\
&- \left[1 - \gamma_1(1,d)(1 - F(s_1^{s,T_1}))\right] \left[1 - Pr(2 \leq Q^{n,T_2} \leq T_1 | Q^{n,T_2} > 1)\right] \\
&= \left[\gamma_1(1,d) \left(F(s_1^{s,T_1}) - F(s_1^{s,T_2})\right)\right] \left[1 - Pr(2 \leq Q^{n,T_2} \leq T_1 | Q^{n,T_2} > 1)\right] \\
&> \left[\gamma_1(1,d) \left(F(s_1^{s,T_1}) - F(s_1^{s,T_2})\right)\right] \left[1 - Pr(2 \leq Q^{pa,T_2} \leq T_1 | Q^{pa,T_2} > 1)\right] \\
&= \left[\gamma_1(1,d) \left(F(s_1^{s,T_1}) - F(s_1^{s,T_2})\right)\right] \prod_{t=2}^{T_1} F(s_t^{n,T_2}), \tag{A.29}
\end{aligned}$$

As before, let  $\mathbf{d}_t$  denoting the  $1 \times t$  vector  $(d, \dots, d)$ . Let  $Pr(\alpha_t | h_t = \mathbf{d}_{t-1}, \alpha_1, T_i)$  denote the probability that the naive DM is attentive or inattentive in period  $t$ , conditional on: 1) not having completed the task by period  $t$ , 2) facing the deadline  $T_i$ , and 3) whether or not he was attentive in period 1.

Condition (i) in the statement of the proposition implies that  $Pr(\alpha_{T_1} = 1 | \mathbf{d}_{T_1-1}, \alpha_1 = 1, T_2) \leq \gamma_1(1,d)$ . Conditions (ii) and (iii) thus implies that for  $t > T_1$

$$\begin{aligned}
Pr(\alpha_t = 1 | \mathbf{d}_{t-1}, \alpha_1 = 1, T_2) &\leq \lambda Pr(\alpha_{t-1} = 1 | \mathbf{d}_{t-2}, \alpha_1 = 1, T_2) + \lambda \gamma_1(d,1) Pr(\alpha_{t-1} = 0 | \mathbf{d}_{t-1}, \alpha_1 = 1, T_2) \\
&< \lambda Pr(\alpha_{t-1} = 1 | \mathbf{d}_{t-1}, \alpha_1 = 1, T_2) + \lambda \gamma_1(d,1). \tag{A.30}
\end{aligned}$$

Equation (A.30) thus shows that if  $Pr(\alpha_{t-1} = 1 | \mathbf{d}_{t-2}, \alpha_1 = 1, T_2) \leq \gamma_1(d,1)$  and  $\lambda < 1/2$ , then  $Pr(\alpha_t = 1 | \mathbf{d}_{t-1}, \alpha_1 = 1, T_2) < \gamma_1(d,1)$  and  $Pr(\alpha_t = 1 | \mathbf{d}_{t-1}, \alpha_1 = 1, T_2) < 2\lambda \gamma_1(d,1)$ . A simple induction thus shows that if  $\lambda < 1/2$ , then  $Pr(\alpha_t = 1 | \mathbf{d}_{t-1}, \alpha_1 = 1, T_2) < 2\lambda \gamma_1(d,1)$  for all  $t > T_1$ .

Thus the probability of being attentive in any period  $t > T_1$  conditional on not having completed the task by that period is bounded above by  $2\lambda \gamma_1(d,1)$  when  $\lambda < 1/2$ . And overall, this implies that for  $2\lambda(T_2 - T_1) < 1$ , the probability of completing the task after period  $T_1$ , conditional on not completing it by period  $T_2$ , is bounded above by  $2\lambda \gamma_1(d,1)(T_2 - T_1)$ .

Thus by equation (A.29),

$$\begin{aligned} Pr(Q^{n,T_2} \leq T_2) - Pr(Q^{n,T_1} \leq T_1) &= Pr(Q^{n,T_2} \leq T_1) + Pr(T_1 < Q^{n,T_2} \leq T_2 | Q^{n,T_2} > T_1) - Pr(Q^{n,T_1} \leq T_1) \\ &< 2\lambda\gamma_1(d, 1)(T_2 - T_1) - \left[ \gamma_1(1, d) \left( F(s_1^{s,T_1}) - F(s_1^{s,T_2}) \right) \right] \prod_{t=2}^{T_1} F(s_t^{n,T_2}) \end{aligned}$$

is negative when  $\lambda$  is sufficiently small such that

$$2\lambda(T_2 - T_1) < \left( F(s_1^{s,T_1}) - F(s_1^{s,T_2}) \right) \prod_{t=2}^{T_1} F(s_t^{n,T_2}).$$

□

*Proof of Proposition 8, part 1.* Let  $f$  be the density function of  $F$ , and let  $M$  denote the maximum  $f$  on  $[\underline{\zeta}, \bar{\zeta}]$ . By assumption  $\iota > 0$ .

*Step 1.* I will begin with the simpler case in which  $b_t \equiv b$  for all  $t$ . For any  $\epsilon > 0$ , Lemma 6 implies that there is a sufficiently high  $T_\epsilon$  such that  $V_0^{\text{pa}, T} > b + \bar{\zeta} - \epsilon^2$ . This then implies that the perfectly attentive DM obtains a payoff that's at least  $b + \bar{\zeta} - \epsilon$  with probability at least  $1 - \epsilon$ . Thus with probability at least  $1 - \epsilon$ , the DM completes the task by using a threshold strategy  $s_t^s > \bar{\zeta} - \epsilon$ . Let  $t^\dagger$  be such that  $s_t^s > \bar{\zeta} - \epsilon$  whenever  $t \geq t^\dagger$ . Thus

$$\prod_{t=t^\dagger}^T F(s_t^s) < \epsilon$$

or

$$\sum_{t=t^\dagger}^T \log F(s_t^s) < \log(\epsilon). \tag{A.31}$$

Now a Taylor expansion shows that

$$\begin{aligned} \sum_{t=t^\dagger}^T \log F(s_t^s) &= - \sum_{t=t^\dagger}^T \left[ \sum_{i=1}^{\infty} (1 - F(s_t^s))^i / i! \right] \\ &> - \sum_{t=t^\dagger}^T \left[ \sum_{i=1}^{\infty} (1 - F(s_t^s))^i \right] \\ &= - \sum_{t=t^\dagger}^T \frac{1 - F(s_t^s)}{F(s_t^s)} \\ &> - \sum_{t=t^\dagger}^T \frac{1 - F(s_t^s)}{F(\bar{\zeta} - \epsilon)}. \end{aligned} \tag{A.32}$$

Now for this same deadline  $T_\epsilon$ , the probability that a naive DM never completes the task satisfies

$$1 - \Pr(Q^n \leq T_\epsilon) < \prod_{t=t^\dagger}^T [1 - \underline{z}(1 - F(s_t^s))]$$

Taking logs and Taylor expanding,

$$\begin{aligned} \log(1 - \Pr(Q^n \leq T_\epsilon)) &< \sum_{t=t^\dagger}^T \log[1 - \underline{z}(1 - F(s_t^s))] \\ &= -\underline{z} \sum_{t=t^\dagger}^T \left[ \sum_{i=1}^{\infty} (1 - F(s_t^s))^i / i! \right] \\ &< -\underline{z} \sum_{t=t^\dagger}^T (1 - F(s_t^s)) \\ &< -\underline{z} F(\bar{\xi} - \epsilon) \log(\epsilon), \end{aligned} \tag{A.33}$$

where the last line follows from equations (A.31) and (A.32).

But now the expression in Equation (A.33) approaches  $-\infty$  as  $\epsilon \rightarrow 0$ , from which it follows that  $\Pr(Q^n \leq T_\epsilon) \rightarrow 1$  as  $\epsilon \rightarrow 0$ .

*Step 2 (sketch)* I now prove the more general statement in the proposition. Let  $Q_*^n$  denote the stopping time corresponding to the model in which period  $t$  payoffs are  $b^* + \xi_t$ . Step 1 is easily generalized to show that for each  $\epsilon > 0$ , there is a  $\Delta_\epsilon$  such that if  $T_{t^\dagger, \epsilon} = t^\dagger + \Delta_\epsilon$   $\Pr(Q_*^n \leq T_{t^\dagger, \epsilon} | Q_*^n > t^\dagger) > 1 - \epsilon/2$  for all  $t^\dagger$ . The reason is as follows. For the sophisticated DM, equations (A.31) and (A.32) generalize immediately to the  $\Delta_\epsilon$  period subgame. For the naive DM, equation (A.33) also carries over to the  $\Delta_\epsilon$  subgame because all that matters is that the probability of being attentive is bounded from below by  $\underline{z}$ .

Next, it is easy to show that the threshold strategies  $s_t^n$  are continuous in the payoffs  $b_1, \dots, b_T$ . Thus by continuity,  $\Pr(Q_*^n \leq T_{t^\dagger, \epsilon} | Q_*^n > t^\dagger) \rightarrow \Pr(Q^n \leq T_{t^\dagger, \epsilon} | Q^n > t^\dagger)$  as  $t^\dagger \rightarrow \infty$ .

Now choose some  $\epsilon > 0$ . And let  $\Delta_\epsilon$  be the value such that  $\Pr(Q_*^n \leq T_{t^\dagger, \epsilon} | Q_*^n > t^\dagger) > 1 - \epsilon/2$  for all  $t^\dagger$  and  $T_{t^\dagger, \epsilon} = t^\dagger + \Delta_\epsilon$ . Now fixing  $\Delta_\epsilon$ , we can find a high enough  $t^\dagger$  be such that  $\Pr(Q^n \leq T_{t^\dagger, \epsilon} | Q^n > t^\dagger) > \Pr(Q_*^n \leq T_{t^\dagger, \epsilon} | Q_*^n > t^\dagger) - \epsilon/2$ . Combined, this shows that  $\Pr(Q^n \leq T_{t^\dagger, \epsilon} | Q^n > t^\dagger) > 1 - \epsilon$ .  $\square$

*Proof of Proposition 8, part 2.* Let  $\Pr(\alpha_t = 1 | \mathbf{d}_{t-1})$  denote the probability that the naive DM

is attentive in period  $t$ , conditional on not yet having completed the task. Note that

$$\begin{aligned} Pr(\alpha_t = 1 | \mathbf{d}_{t-1}) &= \frac{\gamma_t(1, d) Pr(\alpha_{t-1} = 1 | \mathbf{d}_{t-2}) F(s_{t-1}^n)}{1 - Pr(\alpha_{t-1} = 1 | \mathbf{d}_{t-2}) F(s_{t-1}^n)} \\ &< \gamma_t(1, d) Pr(\alpha_{t-1} = 1 | \mathbf{d}_{t-2}) F(s_{t-1}^n) \\ &\leq \bar{z} Pr(\alpha_{t-1} = 1 | \mathbf{d}_{t-2}) F(s_{t-1}^n). \end{aligned}$$

Thus

$$Pr(\alpha_t = 1 | \mathbf{d}_{t-1}) < \bar{z}^t. \quad (\text{A.34})$$

Now for any  $t^\dagger$  and  $\epsilon > 0$ , Lemma 8 implies that there exists a deadline length  $T_{t^\dagger}$  such that  $Pr(Q^{n, T_{t^\dagger}} \leq t^\dagger) < \epsilon$ . This implies that the probability of ever completing the task is given by

$$\begin{aligned} Pr(Q^{n, T_{t^\dagger}} \leq T_{t^\dagger}) &= \epsilon + \sum_{t=t^\dagger+1}^{T_{t^\dagger}} Pr(\alpha_t = 1 | \mathbf{d}_{t-1}) \\ &< \epsilon + \sum_{t=t^\dagger+1}^{T_{t^\dagger}} \bar{z}^t \\ &< \epsilon + \frac{\bar{z}^{t^\dagger}}{1 - \bar{z}^{t^\dagger}} \end{aligned}$$

But  $\frac{\bar{z}^{t^\dagger}}{1 - \bar{z}^{t^\dagger}} \rightarrow 0$  as  $t^\dagger \rightarrow \infty$ , which shows that  $Pr(Q^{n, T_{t^\dagger}} \leq T_{t^\dagger})$  can be made arbitrarily small with a large enough  $T_{t^\dagger}$   $\square$

*Proof of Proposition 9, part 1. Part 1.* By assumption A2, as  $\kappa_t^1 \rightarrow 1$  for all  $t \leq T_1$ ,  $Pr(Q^{n, T_1} \leq T_1) \rightarrow Pr(Q^{pa, T_1} \leq T_1)$ . Similarly,  $Pr(Q^{n, T_2} \leq T_2 | Q^{n, T_2} > \Delta) \rightarrow Pr(Q^{pa, T_2} \leq T_2 | Q^{pa, T_2} > \Delta)$  as  $\kappa_{\Delta+t}^2 \rightarrow 1$  for all  $t \leq T_1$ . But now the proof of part 2 of Proposition 6 shows that  $Pr(Q^{pa, T_2} \leq T_2 | Q^{pa, T_2} > \Delta) \geq Pr(Q^{pa, T_1} \leq T_1)$ . Thus

$$Pr(Q^{n, T_2} \leq T_2) = Pr(Q^{n, T_2} \leq \Delta) + (1 - Pr(Q^{n, T_2} \leq T_2)) Pr(Q^{n, T_2} \leq T_2 | Q^{n, T_2} > \Delta) > Pr(Q^{n, T_1} \leq T_1)$$

as  $\kappa_t^1 \rightarrow 1$  for all  $t \leq T_1$ .

The proof for sophisticated DMs follows identically.  $\square$

*Proof of Proposition 9, part 2.* To simplify notation, set  $q^i = Pr(Q^{n, T_i} \leq T_i - 1)$ , and set  $p^i =$



$Pr(\alpha_{T_i} = 1 | \mathbf{d}_{T_i-1})$ . That is,  $q^i$  is the probability that a naive DM completes the task by period  $T_i - 1$  when facing the deadline  $T_i$ , whereas  $p^i$  is the probability of being attentive in period  $T_i$ , conditional on not having completed the task by that time.

The effect of adding a strength  $\kappa \equiv \kappa_{T_1}^1 = \kappa_{T_2}^2$  cue is that  $p^i$  is transformed to  $p^i(\kappa) = p^i + (1 - p^i)\kappa$ .

Suppose that

$$\left[ q^1 + (1 - q^1)p^1(\kappa)Pr(b_{T_1} + \xi_{T_1} \geq 0) \right] - \left[ q^2 + (1 - q^2)p^2(\kappa)Pr(b_{T_2} + \xi_{T_2} \geq 0) \right] > 0 \quad (\text{A.35})$$

when  $\kappa = 0$ .

Differentiating the expression in (A.35) with respect  $\kappa$ , and noting that  $(Pr(b_{T_1} + \xi_{T_1} \geq 0) \leq Pr(b_{T_2} + \xi_{T_2} \geq 0))$  by assumption, yields

$$(1 - q^1)(1 - p^1)Pr(b_{T_1} + \xi_{T_1} \geq 0) - (1 - q^2)(1 - p^2)Pr(b_{T_2} + \xi_{T_2} \geq 0) \quad (\text{A.36})$$

$$= (1 - q^2)p^2Pr(b_{T_2} + \xi_{T_2} \geq 0) - (1 - q^1)p^1Pr(b_{T_1} + \xi_{T_1} \geq 0) \quad (\text{A.37})$$

$$- (1 - q^2)Pr(b_{T_2} + \xi_{T_2} \geq 0) + (1 - q^1)Pr(b_{T_1} + \xi_{T_1} \geq 0) \quad (\text{A.38})$$

$$< q^1 - q^2 - (1 - q^2)Pr(b_{T_2} + \xi_{T_2} \geq 0) + (1 - q^1)Pr(b_{T_1} + \xi_{T_1} \geq 0) \quad (\text{A.39})$$

$$< (q^1 - q^2)(1 - Pr(b_{T_2} + \xi_{T_2} \geq 0)) \quad (\text{A.40})$$

$$= 0$$

where equation (A.39) is obtained by substituting the inequality (A.35) into the expression in lines (A.37), (A.38).  $\square$

*Proof of Proposition 9, part 3.* To simplify notation, set  $q_t^i = Pr(Q^{n,T_i} \leq t)$ , and set  $p_t^i = Pr(\alpha_t = 1 | \mathbf{d}_{T_i-1})$ . That is,  $q_t^i$  is the probability that a naive DM completes the task by period  $t$  when facing the deadline  $T_i$ , whereas  $p^i$  is the probability of being attentive in period  $t$ , conditional on not having completed the task by that time. Let  $P_t^i(\alpha_t) = Pr(Q^{n,T_i} \leq T_i | \alpha_t, Q^{n,T_i} > t - 1)$  denote the probability of completing the task conditional on i) facing the deadline  $T_i$ , ii) period  $t$  attentiveness  $\alpha_t$ , and iii) not having completed the task before period  $t$ .

The effect of adding a strength  $\kappa \equiv \kappa_t^1 = \kappa_{t+\Delta}^2$  cue in periods  $t$  and  $t + \Delta$ , respectively, is that  $p_t^1$  is transformed to  $p_t^1(\kappa) = p_t^1 + (1 - p_t^1)\kappa$  and  $p_{t+\Delta}^2$  is transformed to  $p_{t+\Delta}^2(\kappa) = p_{t+\Delta}^2 + (1 - p_{t+\Delta}^2)\kappa$

Suppose that  $Pr(Q^{n,T_1} \leq T_1) - Pr(Q^{n,T_2} \leq T_2) > 0$  when  $\kappa = \kappa_t^1 = \kappa_{t+\Delta}^2 = 0$ . Then, equivalently,

$$[q_{t-1}^1 + (1 - q_{t-1}^1)p_t^1(\kappa)P_t^1(1) + (1 - q_{t-1}^1)(1 - p_t^1(\kappa))P_t^1(0)] \quad (\text{A.41})$$

$$- [q_{t-1+\Delta}^2 + (1 - q_{t-1+\Delta}^2)p_{t+\Delta}^2(\kappa)P_{t+\Delta}^2(1) + (1 - q_{t-1+\Delta}^2)(1 - p_{t+\Delta}^2(\kappa))P_{t+\Delta}^2(0)] > 0 \quad (\text{A.42})$$

when  $\kappa = 0$ . Equivalently,

$$[q_{t-1}^1 + (1 - q_{t-1}^1)p_t^1(\kappa)(P_t^1(1) - P_t^1(0)) + (1 - q_{t-1}^1)P_t^1(0)] \quad (\text{A.43})$$

$$- [q_{t-1+\Delta}^2 + (1 - q_{t-1+\Delta}^2)p_{t+\Delta}^2(\kappa)(P_{t+\Delta}^2(1) - P_{t+\Delta}^2(0)) + (1 - q_{t-1+\Delta}^2)P_{t+\Delta}^2(0)] > 0 \quad (\text{A.44})$$

Differentiating the expression in lines (A.43)-(A.44) with respect to  $\kappa$  yields

$$[(1 - q_{t-1}^1)(1 - p_t^1(\kappa))(P_t^1(1) - P_t^1(0))] - [(1 - q_{t-1+\Delta}^2)(1 - p_{t+\Delta}^2(\kappa))(P_{t+\Delta}^2(1) - P_{t+\Delta}^2(0))] \quad (\text{A.45})$$

By assumption,  $\gamma_t^1 = \gamma_{t+\Delta}^2$  for all  $t \leq T_1$ . Combining this with reasoning similar to that in Part 2 of Proposition 6 shows that Condition  $I(T_1, T_2)$  guarantees that

$$P_{t+\Delta}^2(1) \geq P_t^1(1) \quad (\text{A.46})$$

$$P_{t+\Delta}^2(0) \geq P_t^1(0) \quad (\text{A.47})$$

$$P_{t+\Delta}^2(1) - P_{t+\Delta}^2(0) \geq P_t^1(1) - P_t^1(0) \quad (\text{A.48})$$

The assumption that  $\gamma_t^1 = \gamma_{t+\Delta}^2$  for all  $t \leq T_1$  also implies that  $p_t^1 > p_{t+\Delta}^2$ . Thus, if  $q_{t-1}^1 > q_{t-1+\Delta}^2$ , then, using (A.48), the expression in equation (A.45) is negative.

Otherwise, when  $q_{t-1}^1 < q_{t-1+\Delta}^2$ , combining the inequality in lines (A.43)-(A.44) with

inequality (A.48) shows that

$$\begin{aligned} & [(1 - q_{t-1}^1)(1 - p_t^1(\kappa))(P_t^1(1) - P_t^1(0))] - [(1 - q_{t-1+\Delta}^2)(1 - p_{t+\Delta}^2(\kappa))(P_{t+\Delta}^2(1) - P_{t+\Delta}^2(0))] \\ & < q_{t-1}^1 - q_{t-1+\Delta}^2 + (1 - q_{t-1}^1)P_t^1(1) - (1 - q_{t-1+\Delta}^2)P_{t+\Delta}^2(1) \end{aligned} \quad (\text{A.49})$$

$$\leq q_{t-1}^1 - q_{t-1+\Delta}^2 + (1 - q_{t-1}^1)P_{t+\Delta}^2(1) - (1 - q_{t-1+\Delta}^2)P_{t+\Delta}^2(1) \quad (\text{A.50})$$

$$= (q_{t-1}^1 - q_{t-1+\Delta}^2)(1 - P_{t+\Delta}^2(1)) \quad (\text{A.51})$$

$$< 0. \quad (\text{A.52})$$

In the computations above, simple algebra shows that (A.49) is a consequence of the inequality in lines (A.43)-(A.44), and (A.50) is a consequence of (A.46).  $\square$

### A.4.3 Proofs for rebate market results

*Proof of Proposition 10.* As argued in the text, it is clearly not optimal to create rebates for sophisticated consumers.

Consider now naive decision makers. Suppose that when  $v = v_1$ , the optimal rebate policy is  $p_T^*, r_T^* > 0$ . This implies that  $p_T^* - \theta\mu_T r_T^* > v_1$ , where  $\mu_T$  is the probability that the DM applies for the rebate; otherwise the profit maximizing offer would set  $p_T = v_1$  and  $r_T = 0$ . When  $v$  is raised by some amount  $\delta > 0$  to  $v = v_1 + \delta$ , it still follows that  $p_T^* + \delta - \theta\mu_T r_T^* > v_1 + \delta$ .

Note that a change in  $v$  will not change the DM's redemption probability conditional on purchasing, or the perceived value of the rebate. Thus if a positive rebate is better than no rebate for  $v = v_1$ , simply increasing the upfront price by  $\delta$  when  $v$  is increased to  $v_1 + \delta$  increases the profits by  $\delta$ , and is likewise preferred to no rebate.

I now show that a rebate will be offered for a high enough  $v$ . For all  $r_T > -\underline{\xi}$ , the DM's redemption strategy will not depend on  $r$ . So for  $r_T > -\underline{\xi}$ , let  $v_{e,T}$  denote the DM's expected redemption effort. Note, also, that for  $r_T > -\underline{\xi}$ , the DM's *perceived* probability of applying for the rebate is 1. Then for  $r_T > -\underline{\xi}$ , the DM's expected payoff conditional on offer  $p_T, r_T$  is  $v - p_T + \theta(r - v_{e,T})$ . Profit maximization implies that  $p_T = v + \theta(r - v_{e,T})$ , and the price

floor assumption thus requires that

$$v + \theta(r - v_{e,T}) - r \geq \underline{p},$$

from which it follows that

$$r \leq \frac{v - \theta v_{e,T} - \underline{p}}{1 - \theta} \equiv n_T(v).$$

Note that  $n_T(v)$  grows without bound as  $v$  approaches infinity.

Now let  $\theta\mu_T$  denote the DM's actual probability of applying for the rebate when  $r + \underline{\xi} > 0$ . Note that  $\mu_T$  does not depend on  $r$  in the region  $r > -\underline{\xi}$ . Note, also, that  $\mu_T < 1$  by assumption that the DM is inattentive. Now when  $n_T(v) > -\underline{\xi}$  is high enough and generates profit  $v + \theta(r - v_{e,T}) - r \geq \underline{p}$ , the monopolist's profit from setting  $r_T = n_T(v)$  and  $p_T = v + \theta(r - v_{e,T})$  will be

$$\begin{aligned} \pi_T &= v + \theta r - \theta v_{e,T} - \theta \mu_T r - c \\ &= v + \theta(1 - \mu_T)n_T(v) - \theta v_{e,T} - c. \end{aligned}$$

But for high enough  $n_T(v)$ , it is clear that  $v - c + \theta(1 - \mu_T)n_T(v) - \theta v_{e,T} > v - c$ .  $\square$

*Proof of Proposition 11.* As shown in the proof of Proposition 10, for a high enough  $v = c - L$  the profit from offering a rebate  $r = n_T(v)$  will be given by

$$\theta(1 - \mu_T)n_T(v) - \theta v_{e,T} - L. \tag{A.53}$$

As before,  $\mu_T$  is constant in  $r$  in the region  $r = n_T(v) > -\underline{\xi}$ . Moreover,  $n_T(v) > -\underline{\xi}$  for high enough  $v$ . Plainly, then, expression (A.53) is positive for a high enough  $n_T(v)$ .  $\square$

*Proof of Proposition 12. Part 1* On the one hand, a longer deadline always increase the DM's *perceived* value of the rebate. On the other hand, if a longer deadline increases the redemption probability, then it also increases the cost associated with a rebate.

*Part 2* As show in Proposition 8, for any  $r > 0$ , redemption probability will approach 1 as  $T \rightarrow \infty$ . But profits when  $r > 0$  are

$$v + \theta(1 - \mu_T)r - \theta v_{e,T} - c$$

which approach  $v - c - \theta v_{e,T} < v - c$  as  $\mu_T \rightarrow 1$ . □

#### A.4.4 Proofs for reminder advertising results

*Proof of Proposition 13, part 1.* Suppose that a fraction  $\phi_t$  of consumers are attentive in period  $t$ . When a message is not sent in period  $t + 1$ , rehearsal implies that the fraction of consumers who will be attentive that period will be

$$\phi_{t+1} = \phi_t \psi \tag{A.54}$$

, where  $\psi = \ell g(1, a, 0) + (1 - \ell)g(0, d, 0)$ .

Now since all consumers are initially inattentive, the payoff from sending no messages is 0. Consider now the payoff from sending a single message in period 1, and no other messages. In period 1, a fraction  $w$  of consumers will be attentive. Equation (A.54) then implies that the fraction of consumers attentive in period  $t$  will be  $\phi_t = \psi^{t-1}w$ . The total payoffs, therefore, are

$$\sum_{t=1}^{\infty} \delta_o^t \psi^{t-1} w \ell - \delta c = \delta \frac{w \ell}{1 - \delta_o \psi} - \delta c \tag{A.55}$$

Now the quantity in equation (A.55) is the highest payoff a single message can ever generate: If some consumers would already be attentive in the period that the message is sent, then the payoff from sending the message would be smaller. If messages are to be sent in the future, then the incremental payoff from sending the message in the current period is also diminished. □

*Proof of Proposition 13, part 2.* The fraction of attentive consumers after some number  $n$  of message will be  $1 - (1 - w)^n$ . Thus, the payoff of sending message  $n + 1$  in some period  $t$  will be

$$\sum_{\tau=t}^{\infty} [(1 - w)^n - (1 - w)^{n+1}] \delta_o^\tau - \delta_o^t c \ell = \delta_o^t \left[ \left( (1 - w)^n - (1 - w)^{n+1} \right) \frac{\ell}{1 - \delta_o} - c \right] \tag{A.56}$$

Now  $(1 - w)^n - (1 - w)^{n+1} \rightarrow 0$  as  $n \rightarrow \infty$ , and thus the quantity in equation (A.56) approaches 0 as  $n \rightarrow \infty$ . Thus there exists some  $n^\dagger$  such that sending more than  $n > n^\dagger$

messages is not optimal.

The proof of the claim is completed by noting that the expression in (A.56) is positive if and only if  $(1-w)^n - (1-w)^{n+1} \frac{\ell}{1-\delta_o} - c$ . But in this case, the expression in (A.56) is clearly decreasing in  $t$ : sending a message earlier is always preferred to sending a message later. Thus the  $n^\dagger$  message will be sent in the first  $n^\dagger$  periods.  $\square$

*Proof of Proposition 13, part 3.* Suppose, by way of contradiction, that there is some  $t^\dagger$  such that  $m_{t^\dagger} = 1$  but  $m_t = 0$  for all  $t > t^\dagger$ . Then the fraction of consumers attentive in period  $t > t^\dagger$  will be  $\phi_{t^\dagger} \psi^{t-t^\dagger}$ .

If a message is sent in period  $t > t^\dagger$ , but not in any future periods, then the fraction of consumers attentive in period  $t$  will be  $\phi_{t^\dagger} \psi^{t-t^\dagger} + (1 - \phi_{t^\dagger} \psi^{t-t^\dagger})w$ . Rehearsal implies that the fraction of consumers attentive in period  $\tau > t$  will be

$$[\phi_{t^\dagger} \psi^{t-t^\dagger} + (1 - \phi_{t^\dagger} \psi^{t-t^\dagger})w] \psi^{\tau-t} = \phi_{t^\dagger} \psi^{\tau-t^\dagger} + (1 - \phi_{t^\dagger} \psi^{t-t^\dagger})w \psi^{\tau-t} \quad (\text{A.57})$$

The net benefits of sending a message in period  $t > t^\dagger$  are therefore

$$\sum_{\tau=t}^{\infty} (1 - \phi_{t^\dagger} \psi^{t-t^\dagger})w \ell \psi^{\tau-t} - \delta^t c = \sum_{\tau=t}^{\infty} \delta^t \left[ (1 - \phi_{t^\dagger} \psi^{t-t^\dagger}) \frac{w \ell}{1 - \delta_o} - c \right] \quad (\text{A.58})$$

Since  $(1 - \phi_{t^\dagger} \psi^{t-t^\dagger}) \rightarrow 1$  as  $t \rightarrow \infty$ , it follows that the expression in equation (A.58) is positive for high enough  $t$ .  $\square$

*Proof of Lemma 1.* When  $m_t = 1$  each period, the fraction of attentive consumers in period  $t + 1$  is given by

$$\phi_{t+1} = (1 - \phi_t)w + \phi_t w + \phi_t(1 - w)\psi = w(1 - \phi_t\psi) + \phi_t\psi. \quad (\text{A.59})$$

Since  $\phi_0 = 0$ , simple algebra shows that the sequence  $\{\phi_{t+1}\}$  defined in (A.59) is strictly increasing and converges to  $\phi^* = \frac{w}{1-\psi(1-w)}$ .

Now set  $\phi_t = \phi^* - \epsilon_t$ . By definition,  $\epsilon_t \rightarrow 0$  as  $t \rightarrow \infty$ . Consider the payoffs from choosing  $m_{t+1} = 1$  rather than  $m_{t+1} = 0$ . When in period  $t + 1$ , this increases the fraction of

attentive consumers by

$$\begin{aligned}
\Delta_{t+1} &= [(1 - (\phi^* - \epsilon_t)\psi)w + (\phi^* - \epsilon_t)\psi] - [(\phi^* - \epsilon_t)\psi] \\
&= w - w\psi(\phi^* - \epsilon_t) \\
&= \frac{w(1 - \psi)}{1 - \psi(1 - w)} + w\psi\epsilon_t
\end{aligned} \tag{A.60}$$

Now if  $m_\tau = 1$  for all  $\tau \geq t + 2$ , then the period  $t + 1$  increase translates into an additional  $\Delta_{t+1}(1 - w)\psi$  attentive consumers period  $t + 2$ . Intuitively, this is because of those extra consumers who have been made attentive by the period  $t + 1$  message,  $(1 - w)\psi$  will not receive the period  $t + 2$  message and but stay attentive because of rehearsal. Thus, the period  $t + 1$  message increase the fraction of attentive consumers by  $\Delta_{t+1}(1 - w)\psi$  in period  $t + 2$ . Following this logic, the period  $t + 1$  message increase the fraction of consumers attentive in period  $\tau > t + 1$  by  $\Delta_{t+1}((1 - w)\psi)^{\tau - t + 1}$ .

Thus the payoff from sending the message in period  $t + 1$  is

$$\ell \sum_{\tau=t+1}^{\infty} \Delta_{t+1}((1 - w)\psi)^{\tau - t + 1} \delta_o^\tau - c\delta^{t+1} = \delta^{t+1} \left[ \frac{w(1 - \psi)}{1 - \psi(1 - w)} + w\psi\epsilon_t - c \right] \tag{A.61}$$

Since  $\epsilon_t$  approaches zero as  $t \rightarrow \infty$ , the quantity in equation (A.61) is positive for all  $t$  if and only if  $\frac{w(1 - \psi)}{1 - \psi(1 - w)} \geq c$ . □

*Proof of Lemma 2.* Obvious. □

*Proof of Proposition 14.* When  $\ell = 1/2$  and  $w = 1$ ,

$$\frac{w(1 - \psi)\ell}{(1 - \psi(1 - w))(1 - \delta_o\psi(1 - w))} = \frac{1}{2} (1 - g(1, a, 0)/2 - g(1, d, 0)/2) > (1 - g(1, d, 0))/2.$$

Since the expression above is continuous in  $\ell$  and  $w$  for all  $g(1, a, 0)$ , there is some  $w' < 1$  such that condition (1.11) holds in an open neighborhood of  $\ell = 1/2$  for  $w > w'$ .

When  $\ell = 1$  and  $w = 1$ ,  $\psi$  approaches 1 as  $g(1, a, 0)$  approaches 1, and thus the expression in (1.11) approaches 0 as  $g(1, a, 0)$  approaches 1. Continuity implies that there are some  $\bar{g} < 1$  and  $w'' < 1$  such that condition (1.11) does not hold in an open neighborhood of  $\ell = 1$  when  $g(1, a, 0) > \bar{g}$  and  $w > w''$ .

When  $w > 2c$ , condition (1.10) holds for all  $\ell \geq 1/2$ .

Thus choosing  $g(1, a, 0) > \bar{g}$  and  $w > \max\{w', w'', 2c\}$  guarantees that there exist  $\ell_2$  and  $\ell_3$  such that condition (1.11) holds when  $\ell \in (\ell_2, \ell_3)$ , condition (1.11) does not hold when  $\ell \in (\ell_3, 1)$ , but condition (1.10) does hold when  $\ell \in (\ell_3, 1)$ .

Now fix  $g(1, a, 0) > \bar{g}$  and  $w > \max\{w', w'', 2c\}$ , and let  $\ell_2$  be the highest value of  $\ell$  smaller than  $1/2$  for which condition(1.11) holds. Since

$$\frac{w\ell}{1 - \delta_o\psi} > \frac{w(1 - \psi)\ell}{(1 - \psi(1 - w))(1 - \delta_o\psi(1 - w))}$$

there exists some  $\ell_1 < \ell_2$  such that condition (1.10) holds but condition (1.11) does not hold for  $\ell \in (\ell_1, \ell_2)$ . □

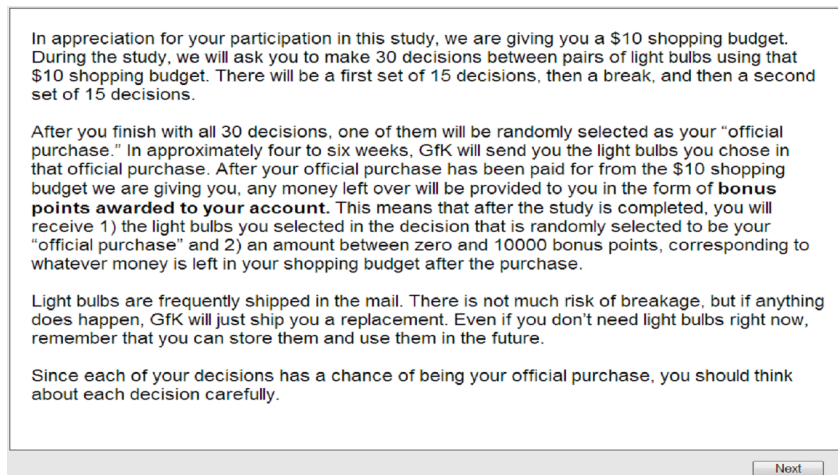


# Appendix B

## Appendix to Chapter 2

### B.1 Details of TESS Experiment

#### Introductory Screen



**Figure B.1:** *Introductory Screen*

We have given you a \$10 shopping budget to purchase a package of light bulbs. Your first 15 purchase decisions will concern the two packages of light bulbs shown below.

**Choice A**  
**Philips 60-Watt-Equivalent  
 Compact Fluorescent Light Bulb, 1-Pack**



**Choice B**  
**Philips 60-Watt Incandescent  
 Light Bulbs, 4-Pack**



[Click for detailed product information](#)

Between the 15 decisions, the only thing that varies is the price. Each of these decisions has a chance of being the one choice (out of 30) that will become your official purchase, so you should think about each purchase carefully. Whatever money you do not spend on the light bulbs, you get to keep: any remaining money will be provided to you as cash-equivalent bonus points. Please think about each decision carefully.

Here is an example of how this might work. After you make all your decisions, suppose that Decision Number 6 from the set below were selected as your official purchase.

- If you had chosen Choice A, you would pay \$2 from your \$10 shopping budget. You would receive the Choice A light bulb package in the mail within 4-6 weeks, as well as the remaining \$10-\$2=\$8 in your shopping budget (You would receive that \$8 in the form of 8000 bonus points credited to your account.)
- If you had chosen Choice B, you would pay \$4 from your \$10 shopping budget. You would receive the Choice B light bulb package in the mail within 4-6 weeks, as well as the remaining \$10-\$4=\$6 in your shopping budget. (You would receive that \$6 in the form of 6000 bonus points credited to your account.)

Now please make your decisions for each of the 15 choices below.

**Figure B.2:** *Baseline Lightbulb Choices (Top of Screen)*

*Detailed Product Information*

Choice:	A	B
Manufacturer:	Philips	Philips
Type:	Compact Fluorescent (CFL)	Incandescent
Number of Bulbs:	1	4
Light Output:	60 Watt-equivalent	60 Watts
Light Output:	900 Lumens	840 Lumens
Color Temperature:	2700K	2700K
Energy Use:	13 Watts	60 Watts
Manufacturer's Home Country:	USA	USA

[Back to Questionnaire](#)

**Figure B.3:** *Detailed Product Information*

Now please make your decisions for each of the 15 choices below.

Decision Number	Choice A 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack	Choice B 60-Watt Incandescent Light Bulbs, 4-Pack
1)	Purchase Choice A for free	Purchase Choice B for \$10
2)	Purchase Choice A for free	Purchase Choice B for \$8
3)	Purchase Choice A for free	Purchase Choice B for \$6
4)	Purchase Choice A for free	Purchase Choice B for \$4
5)	Purchase Choice A for \$1	Purchase Choice B for \$4
6)	Purchase Choice A for \$2	Purchase Choice B for \$4
7)	Purchase Choice A for \$3	Purchase Choice B for \$4
8)	Purchase Choice A for \$4	Purchase Choice B for \$4

Figure B.4: Baseline Product Information (Bottom of Screen)

For this next part of the study, you will have the opportunity to learn more about compact fluorescent light bulbs (CFLs) and incandescent light bulbs. We will focus on the following two issues:

- Total Costs
- Disposal and Warm-Up Time

The discussion of each issue will be followed by a one-question quiz. Please pay close attention to the discussion so that you can correctly answer the quiz question.

Next

Figure B.5: Total Cost Information Screen

For this next part of the study, you will have the opportunity to learn more about compact fluorescent light bulbs (CFLs) and incandescent light bulbs. We will focus on the following two issues:

- Total Costs
- Disposal and Warm-Up Time

The discussion of each issue will be followed by a one-question quiz. Please pay close attention to the discussion so that you can correctly answer the quiz question.

Next

Figure B.6: Disposal and Warm-Up Information Screen

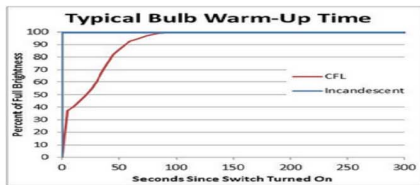
After they burn out, CFLs need proper disposal:

- Because CFLs contain mercury, it is recommended that they be properly recycled, and not simply disposed of in regular household trash. CFLs can be recycled through:
  - Local waste collection sites
  - Mail-back services that you can find online
  - Local retailers, including Ace Hardware, IKEA, Home Depot, and Lowe's, as well as other retailers.
- No special precautions need to be taken to dispose of an incandescent light bulb. Incandescents can be disposed of in regular household trash.

After the light switch is turned on, CFLs take longer to warm up than incandescents:

- An incandescent reaches full brightness immediately.
- A typical CFL can take 60 to 90 seconds to reach its full brightness.

The graph below illustrates this:



**Question:** About how much longer does it take a typical CFL to reach full brightness, as compared to an incandescent?

Type your answer below.  
 to  seconds

Figure B.7: Disposal and Warm-Up Information Screen

For this next part of the study, you will have the opportunity to learn more about light bulbs. We will focus on the following two issues:

1. Number of Bulbs by Sector
2. Sales Trends

The discussion of each issue will be followed by a one-question quiz. Please pay close attention to the discussion so that you can correctly answer the quiz question.

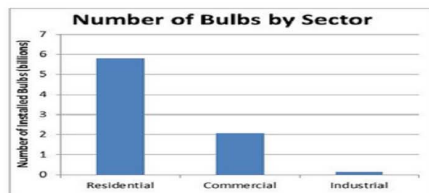
**Figure B.8:** *Control Introductory Screen*

According to official estimates, there are slightly more than eight billion light bulbs installed in the United States.

The US economy can be divided into three major sectors: residential, commercial, and industrial. Each sector has a different number of light bulbs:

- There are about 5.8 billion light bulbs installed in residential buildings in the U.S.
- There are about 2.1 billion light bulbs installed in commercial buildings in the U.S.
- There are about 0.14 billion light bulbs installed in industrial buildings in the U.S.

The graph below illustrates this:



**Question:** About how many more light bulbs are installed in residential buildings compared to commercial buildings in the U.S.?

To answer this question, you can enter whole numbers and/or decimals.

Type your answer below.

billion

**Figure B.9:** *Number of Bulbs by Sector Information Screen*

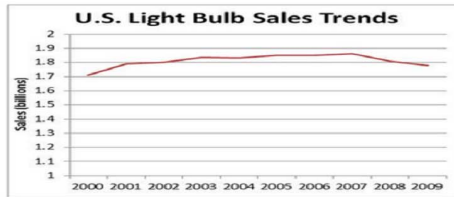
According to official sales data, sales of light bulbs in the United States have had the following trend:

- Sales increased in each year between 2000 and 2007.
- Sales decreased slightly in 2008 and 2009.

Total light bulb sales were different at the end of the decade compared to the beginning:

- Sales in 2000 were just over 1.7 billion bulbs.
- Sales in 2009 were just under 1.8 billion bulbs.

The graph below illustrates this:



**Question:** About how many light bulbs were sold in the United States in 2009?

To answer this question, you can enter whole numbers and/or decimals.

Type your answer below.

billion

**Figure B.10:** Sales Trends Information Screen

Remember, we have given you a \$10 shopping budget to purchase a package of light bulbs. Your second 15 purchase decisions will concern the two packages of light bulbs shown below.

<p><b>Choice A</b> <b>Philips 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack</b></p> 	<p><b>Choice B</b> <b>Philips 60-Watt Incandescent Light Bulbs, 4-Pack</b></p> 
---	---

[Click for detailed product information](#)

Between the 15 decisions, the only thing that varies is the price. Each of these decisions has a chance of being the one choice (out of 30) that will become your official purchase, so you should think about each purchase carefully. Whatever money you do not spend on the light bulbs, you get to keep: any remaining money will be provided to you as cash-equivalent bonus points. Please think about each decision carefully.

Here is an example of how this might work. After you make all your decisions, suppose that Decision Number 21 from the set below were selected as your official purchase.

- If you had chosen Choice A, you would pay \$2 from your \$10 shopping budget. You would receive the Choice A light bulb package in the mail within 4-6 weeks, as well as the remaining \$10-\$2=\$8 in your shopping budget (You would receive that \$8 in the form of 8000 bonus points credited to your account)
- If you had chosen Choice B, you would pay \$4 from your \$10 shopping budget. You would receive the Choice B light bulb package in the mail within 4-6 weeks, as well as the remaining \$10-\$4=\$6 in your shopping budget (You would receive that \$6 in the form of 6000 bonus points credited to your account.)

Now please make your decisions for each of the 15 choices below.

**Figure B.11:** Endline Lightbulb Choices (Top of Screen)

Now please make your decisions for each of the 15 choices below.

Decision Number	Choice A 60-Watt-Equivalent Compact Fluorescent Light Bulb, 1-Pack	Choice B 60-Watt Incandescent Light Bulbs, 4-Pack
16)	Purchase Choice A for free <input type="radio"/>	Purchase Choice B for \$10 <input type="radio"/>
17)	Purchase Choice A for free <input type="radio"/>	Purchase Choice B for \$8 <input type="radio"/>
18)	Purchase Choice A for free <input type="radio"/>	Purchase Choice B for \$6 <input type="radio"/>
19)	Purchase Choice A for free <input type="radio"/>	Purchase Choice B for \$4 <input type="radio"/>
20)	Purchase Choice A for \$1 <input type="radio"/>	Purchase Choice B for \$4 <input type="radio"/>
21)	Purchase Choice A for \$2 <input type="radio"/>	Purchase Choice B for \$4 <input type="radio"/>
22)	Purchase Choice A for \$3 <input type="radio"/>	Purchase Choice B for \$4 <input type="radio"/>
23)	Purchase Choice A for \$4 <input type="radio"/>	Purchase Choice B for \$4 <input type="radio"/>
	Purchase Choice A for \$4 <input type="radio"/>	Purchase Choice B for \$3 <input type="radio"/>

**Figure B.12:** Endline Lightbulb Choices (Bottom of Screen)

## Post-Experiment Survey Questions

**Question 1.** *How important were the following factors in your purchase decision? [Rate from 1-10]*

1. *Energy use*
2. *Time required for the bulb to reach full brightness after it is turned on*
3. *Bulb lifetime*
4. *Mercury content and protocols for proper disposal*
5. *Purchase Price*

**Question 2.** *Do you think that the intent of the study was to ...*

*Select all answers that apply*

1. *Understand the effect of price changes on purchasing patterns*
2. *Measure whether people make consistent purchases in similar situations*
3. *Understand why people buy incandescents vs. CFLs*
4. *Test how well people are able to quantify energy costs*
5. *Test whether ability to quantify energy costs affects purchases of incandescents vs. CFLs*
6. *Test whether the number of bulbs in a package affects purchasing patterns*
7. *Test whether consumer education affects purchases of incandescents vs. CFLs*
8. *Understand what features of lightbulbs are most important to people*
9. *Predict the future popularity of incandescents vs. CFLs*
10. *None of the above*



**Question 3.** Part A: *The typical CFL lasts 8000 hours, or about eight years at typical usage rates. Do you think it costs more or less to buy electricity for that 8000 hours of light from compact fluorescent light bulbs (CFLs) compared to incandescent light bulbs?*

- More
- Less

Part B: *At national average electricity prices, how much [more/less] does it cost to buy electricity for that 8000 hours of light from compact fluorescent light bulbs (CFLs) compared to incandescent light bulbs? Just give your best guess.*

**Question 4.** *Some states and local areas have rebates, low-interest loans, or other incentives available for energy efficiency. These might include rebates for Energy Star appliances or energy efficient light bulbs, low-interest loans for energy-saving home improvements, government-funded weatherization, and other programs. Are any such programs available in your area?*

1. Yes
2. *I think so, but I'm not sure*
3. *I'm not sure at all*
4. *I think not, but I'm not sure*
5. No

**Question 5.** *This question is about hypothetical choices and does not affect your earnings in this study.*

*Suppose that you could get the amount under “Option A” (i.e. \$100), or the amount under “Option B” a year later. Assume it’s no more work for you to receive the money under Option A than under Option B, and that you would receive the money for sure, regardless of when you choose to receive it. Which would you prefer?*

For each row, please choose whether you would prefer the payment under “Option A” or under “Option B”.

	Option A	Option B
1)	\$100 today <input type="radio"/>	\$50 in one year <input type="radio"/>
2)	\$100 today <input type="radio"/>	\$90 in one year <input type="radio"/>
3)	\$100 today <input type="radio"/>	\$100 in one year <input type="radio"/>
4)	\$100 today <input type="radio"/>	\$110 in one year <input type="radio"/>
5)	\$100 today <input type="radio"/>	\$130 in one year <input type="radio"/>
6)	\$100 today <input type="radio"/>	\$150 in one year <input type="radio"/>
7)	\$100 today <input type="radio"/>	\$170 in one year <input type="radio"/>
8)	\$100 today <input type="radio"/>	\$200 in one year <input type="radio"/>
9)	\$100 today <input type="radio"/>	\$250 in one year <input type="radio"/>
10)	\$100 in one year <input type="radio"/>	\$50 in two years <input type="radio"/>
11)	\$100 in one year <input type="radio"/>	\$90 in two years <input type="radio"/>
12)	\$100 in one year <input type="radio"/>	\$100 in two years <input type="radio"/>

Notes: This does not show all of the 18 choices. Participants were randomly assigned to receive either this table or another table that was identical except that the bottom half and top half were switched, so that the one year vs. two year tradeoffs were presented first.

**Question 6.** Please indicate how much you agree or disagree with the following statements:

Select one answer from each row in the grid

[Strongly Agree    Agree    Neutral    Disagree    Strongly Disagree]

1. *It's important to me to fit in with the group I'm with.*
2. *My behavior often depends on how I feel others wish me to behave.*
3. *My powers of intuition are quite good when it comes to understanding others' emotions and motives.*
4. *My behavior is usually an expression of my true inner feelings, attitudes, and beliefs.*
5. *Once I know what the situation calls for, it's easy for me to regulate my actions accordingly.*
6. *I would NOT change my opinions (or the way I do things) in order to please someone else or win their favor.*

## **B.2 Additional TESS Results**

### **Baseline Willingness-to-Pay**

Table A2.1 shows the association between baseline WTP  $v^0$  and a series of individual characteristics. Column 1 shows that men, democrats, and those who report having taken steps to conserve energy have higher demand for CFLs. Columns 2-5 separately test individual variables of environmentalism and political ideology which are correlated, providing additional evidence that liberals tend to have higher WTP. These correlations conform to our intuition and build further confidence that the differences in WTP are meaningful.

The table also provides suggestive evidence on two distortions other than imperfect information and inattention which might justify subsidies and standards. The first is a particular form of agency problem in real estate markets: renters might have lower CFL demand because they might leave the CFLs in the house's light sockets when they move and

be unable to capitalize on their investment. Lacking random or quasi-random assignment in renter vs. homeowner status, Davis (2012) and Gillingham, Harding, and Rapson (2012) correlate durable good ownership with homeowner status conditional on observables. Column 1 replicates their approach in the TESS data, showing no conditional association between WTP and homeowner status. Column 6 shows that the unconditional association is also statistically zero.

The second potential distortion considered in Table A2.1 is present bias. In the post-experiment survey, we estimate the  $\beta$  and  $\delta$  of a quasi-hyperbolic model through a menu of hypothetical intertemporal choices at two different time horizons: \$100 now vs.  $\$m_i^1$  in one year, and \$100 in one year vs.  $\$m_i^2$  in two years. Denoting  $\hat{m}^1$  and  $\hat{m}^2$  as the minimum values at which participant  $i$  prefers money sooner, the long run discount factor is  $\delta_i = 100/\hat{m}^2$ , and the present bias parameter is  $\beta_i = \hat{m}_i^2/\hat{m}_i^1$ . We dropped non-monotonic responses and top-coded  $\hat{m}^1$  and  $\hat{m}^2$  analogously to how we constructed  $v^0$  and  $v^1$ .

If there is a distribution of  $\beta$  and  $\delta$ , consumers with higher  $\beta$  and  $\delta$  should be more likely to purchase CFLs. Column 1 shows that there is a conditional correlation between  $\delta$  and baseline WTP  $v_i^0$ , suggesting that people who are more patient may be more likely to purchase CFLs. However, there is no statistically significant correlation between  $\beta$  and  $v_i^0$ . Column 7 repeats the estimates without any conditioning variables, and the coefficients are comparable. The results in column 1 rule out with 90 percent confidence that a one standard deviation increase in  $\beta$  increases WTP for the CFL by more than \$0.47. In sum, these correlations provide no suggestive evidence in favor of the hypotheses that agency problems or present bias play a role in lightbulb decisions.

**Table B.1:** Association Between Individual Characteristics and CFL Demand

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income (000s)	0.005 (0.006)						
Education (Years)	0.010 (0.148)						
Age	0.003 (0.018)						
Male	0.931 (0.533)*						
Liberal	0.091 (0.389)	0.374 (0.285)					
Party	0.573 (0.344)*		0.562 (0.266)**				
Environmentalist	0.682 (0.791)			1.429 (0.804)*			
Conserve Energy	0.970 (0.525)*				0.863 (0.545)		
Homeowner	0.047 (0.716)					0.116 (0.616)	
Present Bias $\beta$	0.281 (0.298)						0.144 (0.284)
Discount Factor $\delta$	1.215 (0.620)*						0.962 (0.605)
R2	0.03	0.00	0.01	0.00	0.00	0.00	0.01
N	1,163	1,226	1,229	1,221	1,219	1,229	1,178

Notes: Left-hand-side variable is baseline relative WTP for the CFL. \*, \*\*, \*\*\*: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

**Table B.2:** Correlation of Treatment Effects with Self-Monitoring Scale

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Important to fit in	0.206 (0.399)						
Behave as others wish		0.480 (0.384)					
Good intuition for others' motives			0.266 (0.295)				
(-1)*Behavior expresses true feelings				-0.410 (0.332)			
Regulate my actions					-0.218 (0.310)		
(-1)*NOT change opinions to please someone						-0.114 (0.365)	
Self-Monitoring Mean							-0.010 (0.008)
R2	0.57	0.57	0.56	0.57	0.56	0.56	0.57
N	1,185	1,184	1,184	1,184	1,184	1,184	1,188

Notes: This table presents estimates of Equation (2.1) with the addition of Self-Monitoring Scale variables and the interaction of these variables with the treatment indicator. The outcome variable is endline willingness-to-pay for the CFL. Robust standard errors in parenthesis. \*, \*\*, \*\*\*: Statistically significant with 90, 95, and 99 percent confidence, respectively. Observations are weighted for national representativeness.

**Table B.3:** Sensitivity of TESS Welfare Analysis to Assumed Mean Censored Values Assuming Top-Coded and Bottom-Coded WTPs Average \$12 and -\$12, Respectively

(1)	(2)	(3)	(4)	(5)	(6)
CFL Subsidy (\$/package)	Relative WTP ( $v_i^0$ ) of Marginal Consumers (\$/package)	Average Marginal Internality (\$/package)	Demand Density (Share of packages)	Marginal Welfare Effect (\$/package)	Cumulative Welfare Effect (\$/package)
1	0.5	1.92	0.126	0.180	0.180
2	1.5	2.29	0.052	0.041	0.221
3	2.5	2.98	0.028	0.014	0.234
4	3.5	0.69	0.030	-0.085	0.150
6	5	0.69	0.006	-0.027	0.123
8	7	0.69	0.008	-0.050	0.073
10	9	0.69	0.003	-0.022	0.051
$\infty$	12	0.69	0.043	-0.485	-0.434

*Notes:* This table uses the TESS experiment results to calculate the welfare effects at different levels of the CFL subsidy. They replicate Table 8, except with different assumed mean censored values of WTP. Observations are weighted for national representativeness.

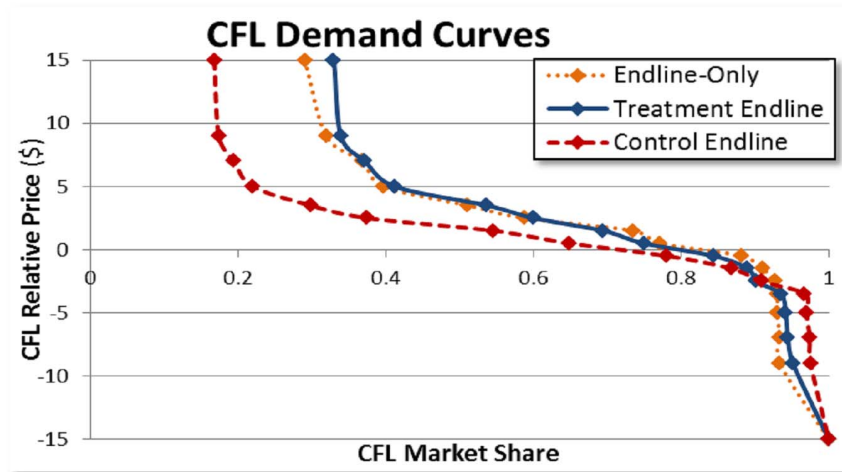
### Sensitivity of TESS Welfare Analysis to Assumed Mean Censored Values

**Table B.4:** Sensitivity of TESS Welfare Analysis to Assumed Mean Censored Values Assuming Top-Coded and Bottom-Coded WTPs Average \$20 and -\$20, Respectively

(1)	(2)	(3)	(4)	(5)	(6)
CFL Subsidy (\$/package)	Relative WTP ( $v_i^0$ ) of Marginal Consumers (\$/package)	Average Marginal Internality (\$/package)	Demand Density (Share of packages)	Marginal Welfare Effect (\$/package)	Cumulative Welfare Effect (\$/package)
1	0.5	2.43	0.126	0.244	0.244
2	1.5	1.95	0.052	0.024	0.268
3	2.5	4.11	0.028	0.046	0.314
4	3.5	0.18	0.030	-0.100	0.213
6	5	0.18	0.006	-0.030	0.183
8	7	0.18	0.008	-0.054	0.129
10	9	0.18	0.003	-0.023	0.106
$\infty$	20	0.18	0.043	-0.850	-0.744

*Notes:* This table uses the TESS experiment results to calculate the welfare effects at different levels of the CFL subsidy. They replicate Table 8, except with different assumed mean censored values of WTP. Observations are weighted for national representativeness.





**Figure B.13:** *Post-Only Treatment Demand Curve*

Notes: Figure B.13 presents the demand curve for the Endline-Only treatment group, along with demand curve for the control group and other treatment group consumers. Observations are weighted for national representativeness.

### B.3 iPad Total Cost Comparison Screen



**Figure B.14:** iPad Total Cost Comparison Screen

Notes: This is the information screen presented to treatment group consumers in the in-store experiment. Numbers in this screen shot represent a consumer buying one CFL at typical purchase prices and national average electricity prices.

## B.4 Appendix to Theoretical Framework

### Proof of Proposition 19

Rewrite welfare as

$$\begin{aligned} W(s) &= C + \sum_k \alpha_k \left[ \int \sigma(v, b_k, c - s)(v - c) dF_{v|b_k}(v|b_k) \right] \\ &= C + \sum_k \alpha_k \left[ \int_{v \geq c - s + b_k} (v - c) dF_{v|b_k}(v|b_k) \right]. \end{aligned} \quad (\text{B.1})$$

Differentiating (B.1) with respect to  $s$  yields

$$\begin{aligned} W'(s) &= \sum_k \alpha_k \left[ -((c - s + b_k) - c) f_{v|b_k}(c - s + b_k|b_k) \right] \\ &= \sum_k \alpha_k (s - b_k) f_{v|b}(c - s + b_k|b_k) \\ &= \sum_k (b_k - s) D'_{b_k}(c - s) \\ &= (B(c - s) - s) D'(c - s). \end{aligned}$$

### Proof of Proposition 20

Suppose that consumers are homogeneous in their bias:  $b_k \equiv b$  for all  $k$ . Then by Proposition 1, an optimal subsidy must satisfy  $s^* = b(c - s^*)$ . We now show that this subsidy attains the first best under homogeneity. Plugging  $s^* = b(c - s^*)$  into the the social welfare function yields

$$\begin{aligned} W(s^*) &= C + \int_{v \geq c - s^* + b(p - s^*)} (v - c) dF(v) \\ &= C + \int_{v \geq c} (v - c) dF(v). \end{aligned} \quad (\text{B.2})$$

But by definition, (B.2) is just the first best level of welfare.

Set  $\psi(s) = s - b(c - s)$ . We must now show that  $\psi(s) = 0$  has a solution  $s^*$ . By definition and by Assumption 1,  $\psi'(s) = 1 + b'(c - s) \geq 1 + \rho$ , where  $\rho > -1$ . This ensures that  $\psi(s)$  has a unique solution.

When consumers are heterogenous in the way specified in the proposition, the argument in the body of the paper shows why the first best can't be obtained.

### Generalizing the Analysis to Partial Internality Reduction

Suppose that consumer bias is given by  $b = b^x + b^y$ , and set

$$B^x(p) = \frac{\sum b^x D'_{b_k}}{D'} \text{ and } B^y(p) = \frac{\sum b^y D'_b}{D'}.$$

Consider now a demand curve  $D^y_{b_k}(p) = 1 - F_{v|b_k}(p + b^y|b_k)$  corresponding to a partial debiasing; namely, elimination of the bias component  $b^x$ . Set  $D^y = \sum_k D^y_{b_k}$ . The reasoning of section 2.5.3 goes through almost verbatim to establish that

$$B^y(p) \approx \frac{D^y(p) - D(p)}{D'(p)}.$$

and

$$B^y(p) \approx \frac{D^y(p) - D(p)}{(D^y)'(p)}.$$

As shown in the paper, the optimal subsidy satisfies  $s^* = B^x + B^y$ . It thus follows that if  $B^x > 0$ —as it would if  $b_x$  corresponds to present-biased undervaluations of the energy costs—then  $\frac{D^y(p) - D(p)}{(D^y)'(p)}$  constitutes an approximate lower bound for the optimal subsidy. Similarly, if  $B^x < 0$ —as it would be if  $b^x$  corresponds to undervaluations of positive attributes of the incandescent that are not debiased by our interventions—then  $\frac{D^y(p) - D(p)}{(D^y)'(p)}$  constitutes an approximate upper bound for the optimal subsidy.

## Appendix C

# Appendix to Chapter 3

### C.1 Additional tables

**Table C.1:** *P-Values of Wilcoxon Rank-Sum tests*

Test	p-value
Average offers in phase 1: PC Market vs. RC Market	$p < 0.01$
First Period MAO's in Phase 2: PC Responders vs. RC Responders	$p = 0.018$
Average MAO's in Phase 2: PC Responders vs. RC Responders	$p = 0.062$
Average MAO's in Phase 2: Matching Group 2 vs. Matching Group 3	$p = 0.38$

**Table C.2:** Minimum Acceptance Thresholds and Cognitive Reflection Test Scores

	First Period				All Periods	
	(1)	(2)	(3)	(4)	(5)	(6)
RC	-13.000** (4.208)	-13.851*** (4.113)	-11.878 (7.585)	-8.573* (3.946)	-9.286** (3.670)	-4.042 (5.098)
CRT		5.317*** (1.222)	6.110* (3.127)		4.453*** (1.054)	6.561** (2.349)
RC*CRT			-1.117 (3.335)			-2.968 (2.573)
Constant	49.000*** (3.051)	39.854*** (4.231)	38.490*** (6.941)	44.373*** (3.624)	36.714*** (3.706)	33.089*** (4.554)
Adj. $R^2$	0.064	0.119	0.107	0.038	0.091	0.095
Observations	75	75	75	1125	1125	1125

OLS Regressions; Clustering by treatment groups; only first period observations are used; Significance levels: \*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$ .

## C.2 Proposer optimization

We can compute the ex-post profit maximizing offer for each period in a matching group, given the observed cumulative distribution function of minimum acceptance thresholds  $F_{mt}(MAO)$  of responders within that matching group:

$$a_{mt}^* = \arg \max_a F_{mt}(a)[k(Y - a)],$$

where  $m$  denotes the individual matching group and  $t$  denotes the time period.

	prof. max. offer	actual avg. offer
PC Matching Group	51.8	56.3
RC Matching Group	45.6	45.8
Mixed Matching Group	39.3	47.5

**Table C.3:** Profit maximizing offers by matching groups

Table C.3 shows the average profit maximizing offer in phase 2 of the experiment conditional on the matching group. Comparing the profit maximizing offers across the different matching groups, it can be seen that the profit maximizing offer in the PC Matching

Group is 6.2 points higher than in the RC Matching Group ( $p = 0.25$ , Ranksum Test based on 10 observations, one per matching group in each treatment), and 12.5 point higher than in the Mixed Matching Group ( $p = 0.08$ ). The profit maximizing offers between the RC matching group and the mixed matching group are not statistically significantly different ( $p = 0.35$ ). Pooling the two matching groups with RC responders and comparing the average profit maximizing offer with the PC matching group which contained PC responders, the profit maximizing offer in the PC matching group is 9.5 points higher ( $p = 0.08$ , Ranksum test based on 15 observations, one for each independent matching group). Moreover, actual average offers are not statistically significantly different from profit maximizing offers in the PC and RC matching groups. The exception is the mixed matching group, in which actual average offers are significantly larger than profit maximizing offers ( $p = 0.08$ , ranksum test based on matching groups). The data therefore again supports the predictions of the theory: Proposers' behavior is close to profit maximizing. Because minimum acceptance thresholds differ depending on responder experience, profit maximizing behavior leads proposers to adopt their offers depending on responder experience.

### C.3 Proofs of Propositions (online publication only)

*Proof of Proposition 17.* We will establish this result under the following more general assumptions: the two proposers have reference points  $r_{P_1}, r_{P_2} \in [0, 1]$ , and the responder has a reference point  $r_R \in [0, 1]$  such that  $r_{P_i} + r_R \leq 1$ .

We begin by showing that both proposers offering  $a = Y$  and the responder accepting one of the offers is, indeed, a SPE. Begin with the responder. His utility from rejecting both offers is  $u_R = 0$ . His utility from accepting one of the offers is  $u_R = Y - \lambda \max(r_R Y - Y, 0) = Y > 0$ , since  $r_R \leq 1$ . Thus it is optimal for the responder to accept. Next, consider the proposers. Both proposers get a financial payoff of 0 when they both offer  $a = Y$ , regardless of whose offer the responder accepts. Now if proposer  $i$  deviates and offers  $a_i < Y$  then the responder will reject his offer with probability 1. This proposer  $i$  will then again end up with a financial payoff of 0. Thus neither proposer has an incentive to deviate.

We now show that there exists no SPE in which both proposers offer  $a_i < Y$  with positive probability. For each of the proposers  $i$ , let  $\alpha_i$  denote their (possibly) mixed strategy. Let  $\underline{a}_i$  denote the lowest offer in the support of  $\alpha_i$ . Let  $H_2$  be the cumulative distribution function corresponding to proposer 2's offers, and set  $\underline{a}'_2 = \inf\{a' : H_2(a') > 0\}$ . We will begin by showing that it is impossible to have  $\underline{a}_1 < \underline{a}'_2 < Y$  in any SPE.

So suppose, by way of contradiction, that  $\underline{a}'_2 < Y$ . Since conditional on accepting an offer, the responder's utility is strictly increasing in the size of the offer, the responder will reject the offer  $\underline{a}_1$  with probability 1. Thus when proposer 1 makes an offer  $\underline{a}_1$ , his utility cannot be greater than zero.

Set

$$a^\dagger := \frac{r_R k Y}{1 - r_R + k r_R}$$

and suppose that proposer 2 makes an offer  $a_i \leq a^\dagger$  with positive probability. Then for proposer 1, an offer of  $\underline{a}_1$  is strictly dominated by an offer of  $a = a^\dagger + \epsilon$ , for a small enough  $\epsilon > 0$ . Notice that if  $a > a^\dagger$ , then  $a > r_R[k(Y - a) + a]$  by definition, and thus the responder derives positive utility from accepting the offer  $a$ . Thus proposer 1's utility from having



offer  $a = a^\dagger + \epsilon$  accepted is

$$\begin{aligned}
u_{P_1} &= k(Y - a) - \lambda \max\{r_{P_1}[k(Y - a) + a] - k(Y - a), 0\} \\
&\geq k(Y - a^\dagger) - k\epsilon - \lambda\epsilon r_R(1 + k) \\
&> 0 \text{ for small enough } \epsilon,
\end{aligned}$$

where in going from the first to the second equation we use the assumption that  $r_{P_1} \leq 1 - r_R$  and the fact that  $a^\dagger = r_R[k(Y - a^\dagger) + a^\dagger]$  by definition. Thus if proposer 2 offers  $a_i \leq a^\dagger$  with positive probability, offering  $a^\dagger + \epsilon$  is strictly better than offering  $\underline{a}_1$  for proposer 1, since the higher offer has a positive probability of generating positive utility to the proposer.

Next, suppose that proposer 2 offers  $a_2 > a^\dagger$  with probability 1. Then with probability 1, the responder would derive positive utility from accepting an offer made by proposer 2. This means that whenever proposer 1 offers  $\underline{a}_1$ , his offer is rejected, while Proposer 2's offer is accepted with probability 1. Now if  $\underline{a}'_2 < Y$ , then proposer 1's utility when he offers  $a = \underline{a}'_2 + \epsilon$  (for a small  $\epsilon$ ) and the offer is accepted is

$$\begin{aligned}
u_{P_1} &= k(Y - a) - \lambda \max\{r_{P_1}[k(Y - a) + a] - k(Y - a), 0\} \\
&= k(Y - \underline{a}'_2) - \lambda \max\{r_{P_1}[k(Y - \underline{a}'_2) + \underline{a}'_2] - k(Y - \underline{a}'_2), 0\} - k\epsilon - \lambda\epsilon r_R(1 + k) \\
&> -\lambda \max\{r_{P_1}[k(Y - \underline{a}'_2) + \underline{a}'_2] - k(Y - \underline{a}'_2), 0\} \text{ for small enough } \epsilon
\end{aligned}$$

Suppose now that  $k \geq 1$  (as in our experiment). If proposer 1 offers  $\underline{a}_1$ , then with probability 1 his utility is at most as good as  $-\lambda r_{P_1}[k(Y - \underline{a}'_2) + \underline{a}'_2]$  (for each possible offer that proposer 2 makes), and thus proposer 1 increases his utility with positive probability when he offers  $\underline{a}_1 + \epsilon$  for some small  $\epsilon > 0$ .

When  $k < 1$  and  $\underline{a}'_2 < Y$ , then proposer 1's utility when he offers  $\underline{a}_1$  is at most as good as  $-\lambda r_{P_1} Y$  with probability 1 (for each possible offer that proposer 2 makes). However, when proposer 1 offers  $a = Y$ , this offer is accepted whenever proposer 2 offers  $a_2 < Y$ ; thus for a positive measure of proposer 2's offers, proposer 1's utility increases from at most  $-\lambda r_{P_1} Y$  to 0 when he offers  $a = Y$  instead of  $a = \underline{a}_1$ . We have thus established that when  $\underline{a}'_2 < Y$ , there is no SPE in which proposer 1 offers  $a < \underline{a}_2$ .

Suppose now that  $\underline{a}_1 = \underline{a}'_2 < Y$ . If  $\underline{a}'_2$  is offered with probability 0 by proposer 2, then a verbatim repetition of the previous arguments establishes a contradiction. So suppose that  $\underline{a}'_2$  is offered with positive probability. In this case, when proposers  $i = 1, 2$  both offer  $\underline{a}'_2$ , at least one of these offers must be rejected with positive probability in the SPE. But then, the proposer whose offer is rejected with positive probability is better off increasing that offer by some tiny  $\epsilon > 0$ , thereby increasing his probability of acceptance by a non-trivial probability, while at the same time decreasing his payoff conditional on acceptance by a negligible amount.

Thus  $\underline{a}_1 \leq \underline{a}'_2 < Y$  cannot be supported in a SPE, and an identical argument shows that  $\underline{a}_2 \leq \underline{a}'_1 < Y$  cannot be supported in a SPE either. This establishes that at least one of the proposers must offer  $a = Y$  with probability 1 in any mixed-strategy equilibrium.

When restricting to pure-strategy equilibria, there cannot be an equilibrium in which proposer 1 offers  $a_1 < Y$  with probability 1, while proposer 2 offers  $a_2 = Y$  with probability 1. A profitable deviation for proposer 2 would be to offer  $a_1 + \epsilon < Y$ . And similar logic shows that proposer 2 can't offer  $a_2 < Y$ .  $\square$

*Proof of Proposition 18.* We prove this result under the more general conditions that the responders have reference points  $r_{R_1} = r_{R_2} = r_R \in [0, 1]$  and the proposer has a reference point  $r_P \leq 1 - r_R$ .

First, notice that regardless of  $j$ 's strategy, responder  $i$  is always willing to accept an offer

$$a \geq \hat{a} := \frac{\lambda r_R k Y}{1 + \lambda(rk - r + 1)}.$$

This is because

$$\hat{a} - \lambda [r_R [k(Y - \hat{a}) + \hat{a}] - \hat{a}] = 0 \tag{C.1}$$

and so receiving an offer  $a \geq \hat{a}$  yields weakly positive utility to responder  $i$ , regardless of what responder  $j$  does. Thus, since an offer  $a > \hat{a}$  is always accepted, the highest offer that can be supported in a SPE is  $\hat{a}$ .

We now show that any offer  $a \leq \hat{a}$  can be supported in a SPE. The candidate equilibrium is one in which 1. the proposer offers  $a$  and both responders choose to accept it with

probability 1, 2. the responders both reject any offer  $a' < a$  and 3. the responders both accept any offer  $a' > a$ . If responder  $i$  deviates from equilibrium and rejects, then his utility is

$$u_R^{deviate} = -\lambda[r_R(k(Y - a) + a)]$$

with probability 1. In contrast, responder  $i$ 's equilibrium payoff is

$$u_R^{equilib} = \frac{1}{2} [a - \lambda[r_R(k(Y - a) + a) - a]] + \frac{1}{2} [-\lambda[r_R(k(Y - a) + a)]] = u_R^{deviate} + \frac{1}{2}(1 + \lambda)a.$$

Thus for  $a > 0$ , the responder is strictly better off accepting, while for  $a = 0$  the responder is weakly better off accepting.

Now clearly, the proposer has no incentive to offer  $a' > a$ . Next, we show that the subgame following an offer of  $a' < a$  has an equilibrium in which both responders reject. To see this, consider the payoff responder  $i$  gets when he deviates. By equation (C.1), accepting an offer  $a' < a \leq \hat{a}$  generates negative utility for responder  $i$ . On the other hand, if responder sticks with his equilibrium strategy and rejects, then all three players get financial payoffs equal to zero, and thus responder  $i$ 's utility is non-negative. Thus responder  $i$  is better off sticking with his equilibrium strategy.

The exact statement of Proposition 18 obtains in the special case that  $r_R = 1/3$ .  $\square$

*Proof of Proposition 19.* The proof of this proposition considers the more general case in which players don't necessarily coordinate on the zero offers equilibrium in the RC market.

By Proposition 17,  $\mu_R^1 = 1$  for a responder from the PC market. By Proposition 18, a responder in the RC market is offered  $a \leq \hat{a} := \frac{\lambda k Y}{3 + (2+k)\lambda}$ . This means that the responder's share of the surplus is at most

$$\bar{\mu}_{RC} := \frac{\hat{a}}{(Y - \hat{a})k + \hat{a}}.$$

And computations reveal that  $\bar{\mu}_{RC} < 1/3$ . Note, however, that it is possible that  $\mu_R^1 = 0 < \bar{\mu}_{RC}$  in case the offer actually went to the other responder in the RC market.

Either way, we have that the responders' period 2 reference points, as a function of  $\gamma$  are

$$\begin{aligned} r_{PC} &= \frac{1 + \gamma}{2} \\ r_{RC} &= \frac{1 - \gamma}{2} + \gamma \mu_{RC}^1. \end{aligned}$$

And since  $1/2 > \mu_{RC}^1$ , it easily follows that  $r_{PC}$  is increasing in  $\gamma$  while  $r_{RC}$  is decreasing in  $\gamma$ .

The statement of the proposition will thus be proven by showing that the minimal acceptable offer is a strictly increasing function of the reference point.

To see this, notice that the lowest offer a responder is willing to accept must satisfy

$$M - \lambda[r[(Y - M)k + M] - M] = 0$$

from which it follows that  $M = \frac{rkY}{1-r+kr}$ , which is an increasing function of  $r$  for any  $k > 0$ .  $\square$

We now formulate a more general version of proposition 20 that does not rely on the assumption that the RC market is characterized by a zero offers equilibrium. In particular, we suppose that in the RC market, responders are offered some  $a \leq \hat{a}$  and that both responders accept the offer in equilibrium. This means that half of the responders receive a share  $\mu_{RC}^1 = 0$  of the surplus, since they do not get to receive the offer, while the other half of the responders receive a share  $\mu_{RC}^1 \geq 0$  of the surplus, since they do get to participate in the offer. We assume that in the ultimatum game, the proposer does not know which of the responders he is facing. That is, the proposer believes that with probability 1/2 he is facing a responder with  $\mu_{RC}^1 = 0$ , and with probability 1/2 he is facing a responder with  $\mu_{RC}^1 \geq 0$ . We let  $M_{RC,L}$  and  $M_{RC,H}$  denote the corresponding MAOs. We now formulate the following more general proposition to characterize the perfect Bayesian equilibria (PBE) of this game:

- Proposition 26.**
1. A PC proposer offers  $M_{PC}$  to a PC responder in any PBE
  2. A PC proposer offers  $M_{RC,H}$  to a RC responder if  $k(Y - M_{RC,H}) > 2k(Y - M_{RC,L})$  and offers  $M_{RC,L}$  otherwise
  3. A RC proposer offers  $M_{RC,H}$  to a RC responder if  $k(Y - M_{RC,H}) > 2k(Y - M_{RC,L})$  and offers

$M_{RC,L}$  otherwise

4. There exists a  $\gamma^\dagger \in (0, 1]$  such that

(a) If  $\gamma \leq \gamma^\dagger$  then a RC proposer always offers  $M_{PC}$  to a PC responder

(b) If  $\gamma \in (\gamma^\dagger, 1)$  then a RC proposer offers  $a < M_{PC}$  to a PC responder

*Proof of Proposition 26. Proof of (1).* Let  $r_{P_{PC}}$  and  $r_{R_{PC}}$  be the period 2 reference points of the proposer and responder who participated in the PC market in phase 1. By Proposition 17, proposers get a zero share of the surplus in the PC market while responders get a full share of the surplus in the PC market. Thus

$$r_{P_{PC}} + r_{R_{PC}} = \left( \frac{1-\gamma}{2} + \gamma \cdot 0 \right) + \left( \frac{1-\gamma}{2} + \gamma \cdot 1 \right) = 1. \quad (C.2)$$

Now conditional on making an offer that the responder will accept in phase 2, the proposer's optimal strategy is to offer  $a = M_{PC}$ . The only thing we have to check is that the proposer's utility from having this offer implemented is non-negative. The proposer's utility from having this offer implemented is:

$$u_P = k(Y - M_{PC}) - \lambda \max\{r_{P_{PC}}[k(Y - M_{PC}) + M_{PC}] - k(Y - M_{PC}), 0\}, \quad (C.3)$$

Equation (C.3) is certainly non-negative if  $r_{P_{PC}}W - k(Y - M_{PC}) \leq 0$ , where  $W = k(Y - M_{PC}) + M_{PC}$ . But by definition,  $M_{PC}$  must satisfy

$$M_{PC} - \lambda \max\{r_{R_{PC}}W - M_{PC}, 0\} = 0 \quad (C.4)$$

from it which it follows that  $r_{R_{PC}}W - M_{PC} > 0$ . Thus

$$r_{P_{PC}}W = (1 - r_{R_{PC}})W < W - M_{PC} = k(Y - M_{PC}), \quad (C.5)$$

from which it follows that equation (C.3) is positive.

*Proof of (2).* Let  $r_{R_{RC}}$  denote the phase 2 reference point of an RC responder. Then because  $r_{R_{RC}} \leq r_{R_{PC}}$ , equation (C.2) implies that  $r_{P_{PC}} + r_{R_{RC}} \leq 1$ . Analogous to equation (C.4) we can establish that  $M_{RC,H} - \lambda \max\{r_{R_{RC,H}}W - M_{RC,H}, 0\} = 0$ , where  $W$  is now defined as

$W = k(Y - M_{RC,H}) + M_{RC,H}$ . Thus analogous to (C.5), we find that  $r_{P_{PC}}W < k(Y - M_{RC,H})$ . This means that in the domain of offers  $a \leq M_{RC,H}$ , the proposer's share of the surplus exceeds his fairness reference point, and thus he acts as a purely profit-maximizing agent in that domain of offers. Statement (2) now follows because the proposer's maximization problem now simply boils down figuring out if he is better off offering  $M_{RC,H}$  and having his offer accepted by all responders or if he is better off offering  $M_{RC,L}$  and having his offer accepted by just half of the responders.

*Proof of (3).* This proof is virtually identical to the proof of (2).

*Proof of (4).* As before, it is clear that conditional on making an offer that the responder will accept in phase 2, the proposer's optimal strategy is to offer  $a = M_{PC}$ . However, when there is no offer that is acceptable to the responder and that will generate non-negative utility to the proposer, the proposer's optimal strategy is to offer  $a < M_{PC}$ . As before, the proposer's utility, as a function of  $\gamma$ , is

$$u_P(\gamma) = k(Y - M_{PC}(\gamma)) - \lambda \max\{r_{P_{RC}}(\gamma)[k(Y - M_{PC}(\gamma)) + M_{PC}(\gamma)] - k(Y - M_{PC}(\gamma)), 0\}. \quad (\text{C.6})$$

But

$$r_{P_{RC}}[k(Y - M_{PC}) + M_{PC}] - k(Y - M_{PC}) = M_{PC}[k(1 - r_{P_{RC}}) + r_{P_{RC}}] + kY(r_{P_{RC}} - 1). \quad (\text{C.7})$$

Now the left-hand side of (C.7) is clearly increasing in  $r_{P_{RC}}$ , while the right-hand side of (C.7) is clearly increasing in  $M_{PC}$ . Thus (C.7) is increasing in  $\gamma$  because  $M_{PC}$  and  $r_{P_{RC}}$  are increasing in  $\gamma$ . From this, it follows that  $u_P(\gamma)$  is decreasing in  $\gamma$ .

Next, we have

$$r_{P_{RC}} + r_{R_{PC}} = \left(\frac{1-\gamma}{2} + \gamma \cdot \mu_{P,PC}\right) + \left(\frac{1-\gamma}{2} + \gamma \cdot 1\right) = 1 + \gamma\mu_{P,RC}, \quad (\text{C.8})$$

where  $\mu_{P,RC} > 0$  is the share of the surplus that the proposer gets in the RC market. As before, we have that  $r_{R_{PC}}W > M_{PC}$ , where  $W = k(Y - M_{PC}) + M_{PC}$ . Thus

$$r_{P_{RC}}W = (1 - r_{R_{PC}} + \gamma\mu_{P,RC})W < W - M_{PC} + \gamma\mu_{P,RC}W = k(Y - M_{PC}) + \gamma\mu_{P,RC}.$$

Thus when  $\gamma = 0$ , we have that  $k(Y - M_{PC}) - r_{P_{RC}}W > 0$ . Moreover, since  $k(Y - M_{PC}) - r_{P_{RC}}W$  is continuous in  $\gamma$ , we know that  $k(Y - M_{PC}) - r_{P_{RC}}W > 0$  for a neighborhood of  $\gamma$  around 0, and thus that  $u_P(\gamma) > 0$  for  $\gamma$  close enough to zero. Moreover, since  $u_P(\gamma)$  is continuous and decreasing in  $\gamma$ , there must exist some  $\gamma^\dagger \in (0, 1]$  such that  $u_P(\gamma)$  is positive if and only if  $\gamma \leq \gamma^\dagger$ .  $\square$

To see that  $\gamma^\dagger$  can sometimes be less than 1 in the above proof, fix  $\gamma$  and let  $\lambda$  be very large, so that both the proposer and responder require approximately a share  $r_{P_{RC}}$  and  $r_{R_{PC}}$ , respectively, of the surplus to derive non-negative utility from the transaction. But since  $r_{P_{RC}} + r_{R_{PC}} > 1$ , there will not be a division of surplus that suits both the proposer and responder.

*Proof of Proposition 21.* We prove that the statement of the proposition holds under more general assumptions about how reference points are formed. In particular, let  $w_0, w_1, \dots$  be an infinite sequence given by  $w_0 = 1$  and  $w_j = \delta^j$  for some  $\delta \in [0, 1]$ . Then let the period  $t$  reference point of player  $i$  be given by

$$r_i^t = (1 - \gamma)(1/2) + \gamma \frac{\sum_{\tau=-T}^{t-1} w_{t-1-\tau} \mu_i^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}}.$$

The definition used in the main body of the paper is obtained as a special case in which  $\delta = 1$ .

*Step 1:* We first show that there is a unique steady state. In any steady state, we must have

$$M^* - \lambda[r_R^*(k(Y - M^*) + M^*) - M^*] = 0, \quad (\text{C.9})$$

which can be rearranged to show that

$$\frac{M^*}{k(Y - M^*) + M^*} = \frac{\lambda r_R^*}{1 + \lambda}. \quad (\text{C.10})$$

As in the proof of Proposition 26, offering  $a^* = M^*$  is clearly optimal for the proposer, conditional on making an offer that the responder will accept. Moreover, since  $r_P^* + r_R^* = 1$

by definition, we can show, analogously to equation (C.5), that

$$r_P^*[k(Y - M^*) + M^*] < k(Y - M^*),$$

from which it follows that the proposer derives positive utility from making an offer  $a^* = M^*$ . Thus the proposer's optimal strategy is to offer  $a^* = M^*$  in any steady state.

Plugging in  $a^* = M^*$  into (C.10), and using the definition of  $r_R^*$ , we now have that

$$r_R^* = (1 - \gamma)(1/2) + \gamma \frac{\lambda}{1 + \lambda} r_R^*. \quad (\text{C.11})$$

Equation (C.11) is a linear equation in  $r_R^*$  with a unique solution given by

$$r_R^* = \frac{(1 - \gamma) + \lambda(1 - \gamma)}{2 + 2\lambda(1 - \gamma)}. \quad (\text{C.12})$$

Thus there can be at most one steady state. We now show that the unique solution does, indeed, correspond to a steady state. First, examination of equation (C.12) shows that  $r_R^* \in (0, 1)$ : since  $(1 - \gamma) < 2$ , it is clear that the numerator is smaller than the denominator. Next, by definition of  $M^*$ , accepting an offer of  $a^* = M^*$  is weakly optimal for the responder. And as we have already established, offering  $a^* = M^*$  is also optimal for the proposer.

*Step 2:* We now show that for each  $\epsilon > 0$ , there exists a  $t \geq 1$  such that  $r_R^t + r_P^t \leq 1 + \epsilon$ . To see this, notice that  $\mu_R^t + \mu_P^t \leq 1$  for  $t \geq 1$ , regardless of the outcome in period  $t$ . Thus

$$\begin{aligned} r_R^t + r_P^t &= (1 - \gamma) + \gamma \left( \frac{\sum_{\tau=-T}^{t-1} w_{t-1-\tau} \mu_R^\tau + w_{t-1-\tau} \mu_P^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \right) \\ &\leq (1 - \gamma) + \gamma \left( \frac{\sum_{\tau=-T}^0 w_{t-1-\tau} \mu_R^\tau + w_{t-1-\tau} \mu_P^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} + \frac{\sum_{\tau=1}^{t-1} w_{t-1-\tau}}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \right) \\ &= 1 + \gamma \left( \frac{\sum_{\tau=-T}^0 w_{t-1-\tau} \mu_R^\tau + w_{t-1-\tau} \mu_P^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} - \frac{\sum_{\tau=-T}^0 w_{t-1-\tau}}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \right) \end{aligned}$$

But

$$\frac{\sum_{\tau=-T}^0 w_{t-1-\tau} \mu_R^\tau + w_{t-1-\tau} \mu_P^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \leq \frac{\sum_{\tau=-T}^0 2w_{t-1-\tau}}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}}$$

and

$$\frac{\sum_{\tau=-T}^0 w_{t-1-\tau}}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \rightarrow 0$$



as  $t \rightarrow \infty$ . Thus for each  $\epsilon > 0$ , there exists a  $t \geq 1$  such that  $r_R^t + r_P^t \leq 1 + \epsilon$ .

*Step 3:* We now show that there is some  $t^\dagger \geq 1$  such that  $a^t = M^t$  for all  $t \geq t^\dagger$ ; that is, for all  $t \geq t^\dagger$ , the proposer derives positive utility from offering  $M^t$  and having that offer accepted.

Set  $r_P^t = 1 - r_R^t + \epsilon^t$ . As in the proof of Proposition 26, we have that  $r_R^t[k(Y - M^t) + M^t] > M^t$ . Thus

$$\begin{aligned} r_P^t[k(Y - M^t) + M^t] &= (1 - r_R^t + \epsilon^t)[k(Y - M^t) + M^t] \\ &< [k(Y - M^t) + M^t] - M^t + \epsilon^t[k(Y - M^t) + M^t] \\ &= k(Y - M^t) + \epsilon^t[k(Y - M^t) + M^t]. \end{aligned}$$

This means that the proposer's utility from offering  $M^t$  is such that

$$u_P^t \geq k(Y - M^t) - \lambda \max(\epsilon^t, 0).$$

Moreover, because  $r_R^t \leq (1 - \gamma)/2 + \gamma = (1 + \gamma)/2$ , it easily follows that

$$M^t = \frac{k\lambda r_R^t Y}{1 + \lambda(1 - r_R^t) + k\lambda r_R^t}$$

is bounded away from  $Y$  (for all possible  $\lambda$ ) as long as  $\gamma < 1$ . Thus we have that for all  $t$ , there is some  $c > 0$  such that  $k(Y - M^t) \geq c$ . By step 2, there is a  $t^\dagger$  such that  $\lambda\epsilon^t < c$  for all  $t \geq t^\dagger$ . Thus there is a  $t^\dagger$  such that  $k(Y - M^t) - \lambda \max(\epsilon^t, 0) > 0$  for all  $t \geq t^\dagger$ .

*Step 4:* We now strengthen step 2 to show that  $|r_P^t + r_R^t - 1| \rightarrow 0$ . By step 3, we now have that  $\mu_R^t + \mu_P^t = 1$  for all  $t \geq t^\dagger$ . Thus for  $t > t^\dagger$ ,

$$\begin{aligned} r_R^t + r_P^t &= (1 - \gamma) + \gamma \left( \frac{\sum_{\tau=-T}^{t-1} w_{t-1-\tau} \mu_R^\tau + w_{t-1-\tau} \mu_P^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \right) \\ &= (1 - \gamma) + \gamma \left( \frac{\sum_{\tau=-T}^{t^\dagger-1} w_{t-1-\tau} \mu_R^\tau + w_{t-1-\tau} \mu_P^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} + \frac{\sum_{\tau=t^\dagger}^{t-1} w_{t-1-\tau}}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \right) \\ &= 1 + \gamma \left( \frac{\sum_{\tau=-T}^{t^\dagger-1} w_{t-1-\tau} \mu_R^\tau + w_{t-1-\tau} \mu_P^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} - \frac{\sum_{\tau=-T}^{t^\dagger-1} w_{t-1-\tau}}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \right) \end{aligned}$$

But since

$$\frac{\sum_{\tau=-T}^{t^\dagger-1} w_{t-1-\tau} \mu_R^\tau + w_{t-1-\tau} \mu_P^\tau}{\sum_{\tau=-T}^{t^\dagger-1} w_{t-1-\tau}} \rightarrow 0$$

and

$$\frac{\sum_{\tau=-T}^{t^\dagger-1} w_{t-1-\tau}}{\sum_{\tau=-T}^{t^\dagger-1} w_{t-1-\tau}} \rightarrow 0$$

as  $t \rightarrow \infty$ , it follows that  $r_R^t + r_P^t \rightarrow 1$  as  $t \rightarrow \infty$ .

*Step 5:* We now finish off the proof of the proposition by proving that the steady state identified in Step 1 is globally stable.

Define  $v_R^t = \frac{\sum_{\tau=-T}^{t-1} w_{t-1-\tau} \mu_i^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}}$ . Define the map  $\zeta : \mathbb{R} \rightarrow \mathbb{R}$  as follows:

$$\zeta(v) = (1 - \gamma)/2 + \gamma v.$$

Define the map  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  as follows:

$$\psi(v) = \frac{\lambda \zeta(v)}{1 + \lambda}.$$

Notice that  $\psi$  is linear in  $v$  and has slope  $\gamma\lambda/(1 + \lambda) < 1$ ; thus  $\psi$  is a contraction and has a unique fixed point. In a steady state,  $r_R^* = \zeta(v_R^*)$ , and thus equation (C.10) implies that

$$\frac{M^*}{k(Y - M^*) + M^*} = \psi(v_R^*). \quad (\text{C.13})$$

But since  $v_R^* = \frac{M^*}{k(Y - M^*) + M^*}$  by definition, it follows that the unique fixed point of  $\psi$  corresponds to the unique steady state.

Now for  $t^\dagger$  defined as in step 3,  $r^t = \zeta(v^t)$  and  $\frac{M^t}{k(Y - M^t) + M^t} = \psi(v_R^t)$  for all  $t \geq t^\dagger$ . Because  $\zeta$  is strictly increasing, each value of  $v_R^t$  corresponds to a unique value of  $M^t$ . Because  $\psi$  is strictly increasing and because  $\frac{M^t}{k(Y - M^t) + M^t}$  is strictly increasing in  $r_R^t$ , each value of  $v_R^t$  also corresponds to a unique value of  $M^t$ . Because  $\zeta$  and  $\psi$  are both continuous functions of  $v$ , showing that  $v_R^t \rightarrow v_R^*$  will thus imply that  $M^t \rightarrow M^*$  and  $r_R^t \rightarrow r_R^*$ . Moreover, since  $|r_R^t + r_P^t - 1| \rightarrow 0$  by Step 4, convergence of  $r_R^t$  will also imply convergence of  $r_P^t$ . And finally, since Step 3 shows that  $a^t = M^t$  for all  $t \geq t^\dagger$ ,  $v_R^t \rightarrow v_R^*$  will thus also imply that  $a^t \rightarrow a^*$ .

Because  $\psi$  is an increasing and linear function of  $v^t$  that crosses the 45-degree line exactly once, it thus follows that  $\psi(v) \in (v^*, v)$  for  $v > v^*$  and  $\psi(v) \in (v, v^*)$  for  $v < v^*$ . By

definition,

$$v_R^{t+1} = \frac{w_0}{\sum_{\tau=-T}^t w_{t-\tau}} \mu_R^t + \left(1 - \frac{w_0}{\sum_{\tau=-T}^t w_{t-\tau}}\right) v_R^t \quad (\text{C.14})$$

is a convex combination of  $v_R^t$  and  $\psi(v_R^t) = \frac{M^t}{k(Y-M^t)+M^t} = \mu_R^t$ , which implies that  $v_R^{t+1} \in (v_R^*, v_R^t)$  if  $v_R^t > v_R^*$ . Similarly, it follows that  $v_R^{t+1} \in (v_R^t, v_R^*)$  if  $v_R^t < v_R^*$ .

For  $t^\dagger$  defined as in step 3, a simple induction thus implies that if  $v_R^{t^\dagger} < v_R^*$ , then  $v_R^t$  will be strictly increasing for  $t \geq t^\dagger$  and bounded from above by  $v^*$ . Similarly, if  $v_R^{t^\dagger} > v^*$ , then  $v_R^t$  will be strictly decreasing for  $t \geq t^\dagger$  and bounded from below by  $v_R^*$ . Because any monotonic and bounded sequence converges,  $v_R^t$  must converge to some  $v^{**} \in [0, 1]$ . Because each value of  $v_R^t$  corresponds to a unique value of  $M^t$ , and because  $\psi$  is continuous in  $v$ , there must, therefore, exist some  $M^{**}$  such that  $M^t \rightarrow M^{**}$ . Thus

$$\lim_{t \rightarrow \infty} \mu_R^t = \lim_{t \rightarrow \infty} \frac{M^t}{k(Y - M^t) + M^t} = \frac{M^{**}}{k(Y - M^{**}) + M^{**}}.$$

It is then easy to show that

$$v^t = \frac{\sum_{\tau=-T}^{t-1} w_{t-1-\tau} \mu_i^\tau}{\sum_{\tau=-T}^{t-1} w_{t-1-\tau}} \rightarrow \frac{M^{**}}{k(Y - M^{**}) + M^{**}}.$$

On the other hand,

$$\psi(v_R^t) = \frac{M^t}{k(Y - M^t) + M^t} \rightarrow \frac{M^{**}}{k(Y - M^{**}) + M^{**}}.$$

But since  $\psi$  is continuous, we therefore have that  $\psi(v^{**}) = v^{**}$ . And because  $\psi$  has a unique fixed point, it must be that  $v^{**} = v_R^*$ , thus completing the proof.  $\square$

*Proof of Proposition 22.* Since  $r_P^t + r_R^t \leq 1$  for all  $t \geq 1$ , the reasoning of Step 3 in the proof of Proposition 21 implies that the proposer will offer  $a^t = M^t$  in all periods  $t \geq 1$ . Thus for  $t \geq 1$ ,  $r^t = \zeta(v_R^t)$  and  $\frac{M^t}{k(Y-M^t)+M^t} = \psi(v_R^t)$ .

As in the proof of Proposition 21, a simple induction thus implies that if  $v_R^1 < v_R^*$ , then  $v_R^t$  will be strictly increasing for  $t \geq 1$  and bounded from above by  $v^*$ . Similarly, if  $v^1 > v_R^*$ , then  $v_R^t$  will be strictly decreasing for  $t \geq 1$  and bounded from below by  $v_R^*$ . But since  $M^t$  is a monotonic function  $\zeta(\cdot)$  of  $v_R^t$  such that  $M^* = \zeta(v_R^*)$ , the result follows.  $\square$