



## Essays in Applied Microeconomics

### Citation

Sándor, László. 2015. Essays in Applied Microeconomics. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

### Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:14226075>

### 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 Applied Microeconomics

A dissertation presented

by

László Sándor

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

Harvard University

Cambridge, Massachusetts

December 2014

© 2014 László Sándor

All rights reserved.

*Dissertation Advisor:*  
**Professor Raj Chetty**

*Author:*  
**László Sándor**

## **Essays in Applied Microeconomics**

### **Abstract**

This dissertation collects three pieces of work. The first chapter documents empirically how Danish households substituted between insurance and liquidity, namely how the up-take of unemployment insurance fell when credit suddenly became more cheaply available for some. The second chapter presents results from a natural field experiment comparing financial and non-financial incentives to promote pro-social behavior. Finally, the third chapter presents the theoretical motivation for and results from a laboratory experiment conducted in Iceland on measuring time preferences conditional on incomes not changing, or correcting for the change when they do.

# Contents

Abstract . . . . .	iii
Acknowledgments . . . . .	ix
<b>1 Does Liquidity Substitute for Unemployment Insurance? Evidence from the Introduction of Home Equity Loans in Denmark</b>	<b>3</b>
1.1 Introduction . . . . .	3
1.2 Institutional Details and Empirical Strategy . . . . .	7
1.2.1 The Danish Unemployment Insurance System . . . . .	7
1.2.2 The Danish Mortgage System and the 1992 Reform . . . . .	12
1.2.3 Econometric Methods . . . . .	15
1.3 Data and Summary Statistics . . . . .	18
1.4 Results . . . . .	24
1.4.1 Main Results . . . . .	24
1.4.2 Placebo Tests and Robustness Checks . . . . .	32
1.4.3 Heterogenous Effects . . . . .	37
1.5 Conclusion . . . . .	39
<b>2 What Policies Increase Prosocial Behavior? An Experiment with Referees at the <i>Journal of Public Economics</i></b>	<b>40</b>
2.1 Introduction . . . . .	40
2.2 Experimental Design . . . . .	42
2.3 Four Sets of Outcomes . . . . .	46
2.3.1 Outcome 1: Acceptance of Referee Invitation . . . . .	46
2.4 Outcome 2: Review Time . . . . .	47
2.4.1 Outcome 3: Review Quality . . . . .	57
2.4.2 Outcome 4: Spillover Effects on Other Journals . . . . .	58
2.5 Lessons for the Peer Review Process . . . . .	60
2.6 Lessons for Increasing Prosocial Behavior . . . . .	62

<b>3</b>	<b>Compensated Discount Functions:</b>	
	<b>An Experiment on the Influence of Expected Income on Time Preference</b>	<b>63</b>
3.1	Introduction . . . . .	63
3.2	Theoretical background . . . . .	69
3.2.1	Overview . . . . .	69
3.2.2	Model . . . . .	71
3.2.3	Deriving $D$ from $\tilde{\lambda}$ . . . . .	73
3.2.4	Discussion: Interpretations and Assumptions . . . . .	75
3.3	Experimental Design . . . . .	76
3.4	Experimental Procedures and Background . . . . .	79
3.4.1	Experimental Sessions . . . . .	79
3.4.2	Payment Process . . . . .	80
3.4.3	Income Verification . . . . .	80
3.4.4	Economic Situation Around the Time of the Experiment . . . . .	81
3.5	Computation of Discount Factors and Statistical Analyses . . . . .	82
3.6	Experimental Results . . . . .	86
3.6.1	Uncompensated Discount Factors . . . . .	86
3.6.2	Compensated Discount Factors . . . . .	90
3.7	Conclusion . . . . .	93
	<b>References</b>	<b>95</b>
	<b>Appendix A Appendix to Chapter 1</b>	<b>100</b>
A.1	Supplementary Tables . . . . .	100
	<b>Appendix B Appendix to Chapter 2</b>	<b>106</b>
B.1	Supplementary Tables . . . . .	106
B.2	Supplementary Figures . . . . .	111
B.3	Data Sources and Variable Definitions . . . . .	115
B.4	Reweighting Methodology . . . . .	116
B.5	Hazard Model Estimates of Treatment Effects on Review Times . . . . .	117
B.6	List of Other Journals Used to Assess Spillover Effects . . . . .	119
B.7	Invitation Emails . . . . .	120
B.8	Reminder and Thank-You Emails . . . . .	123
B.9	Summary of Appendix Tables and Figures . . . . .	126
	<b>Appendix C Appendix to Chapter 3</b>	<b>127</b>
C.1	Proof of Proposition . . . . .	127
C.2	Supplementary Tables . . . . .	129

## List of Tables

1.1	Summary Statistics for the Estimation Sample in 1991 by Liquidity Shock Quartiles, Compared to the Danish Population of the Same Birth Cohorts . . .	22
1.2	Impact of 1992 Mortgage Reform on Unemployment Insurance Participation (treatment with dosage) . . . . .	29
1.3	Impact of 1992 Mortgage Reform on Unemployment Insurance Participation (discrete treatment) . . . . .	33
1.4	Impact of 1992 Mortgage Reform on Unemployment Insurance Participation among Households Stable around 1991 . . . . .	36
1.5	Impact of 1992 Mortgage Reform on Unemployment Insurance Participation by Unemployment Risk Quintiles . . . . .	38
2.1	Description of Treatment Groups . . . . .	43
2.2	Fraction of Referees Who Accept Review Invitation by Treatment Group . . .	47
2.3	Measures of Review Quality by Treatment Group . . . . .	59
3.1	Description of the Sample . . . . .	88
3.2	Parameters by Clusters . . . . .	91
3.3	p-values from Hypothesis Tests . . . . .	94
A.1	Sample Selection . . . . .	101
A.2	Cohorts Affected by Early Retirement Reforms of the Unemployment Insurance System . . . . .	102
A.3	Impact of 1992 Mortgage Reform on Unemployment Insurance Participation among Birth Cohorts 1960-1962 for Years 1987-1994 . . . . .	103
A.4	Supplemental Security Income . . . . .	104
A.5	Unemployment Benefits and Basic Membership Fees . . . . .	105
B.1	Summary Statistics for Experimental Sample . . . . .	106
B.2	Randomization and Selection Tests . . . . .	107
B.3	Median Review Times by Treatment Group . . . . .	108
B.4	Cox Hazard Model Estimates of Treatment Effects on Review Times . . . . .	109
B.5	Spillover Effects on Other Journals . . . . .	110

C.1 Robustness to Selection on Observables and Unobservables . . . . . 130

## List of Figures

1.1	Expected Net Benefits of Unemployment Insurance in the Estimation Sample	10
1.2	Economic Environment	14
1.3	Insurance Up-Take by Treatment Group around the Reform	25
1.4	Impact of 1991 Home Equity on Unemployment Insurance Sign-Up (with controls)	27
1.5	Placebo Tests: Impact of Pre-1991 Home Equity on Unemployment Insurance Sign-Up (with controls)	35
2.1	Pre-Experiment Review Times for Referees who Accept Invitations During Experiment	49
2.2	Review Times by Treatment Group During Experiment	51
2.3	Review Times Before vs. After End of Cash Reward	55
2.4	Heterogeneity in Treatment Effects by Tenure Status	57
B.1	Timeline of Interventions and Outcomes	111
B.2	Review Times by Treatment Group in Extended Sample	112
B.3	Social Incentives and Tenured vs. Untenured Referees	113
B.4	Spillover Effects: Review Times at Other Journals	114
B.5	Control and Four-Week Invitation Emails	121
B.6	Cash and Social Invitation Emails	122
B.7	Control and Four-Week Reminder Emails	123
B.8	Cash and Social Reminder Emails	124
B.9	Thank-You Emails	125

## Acknowledgments

I am grateful to Raj Chetty, Larry Katz, and David Laibson for careful guidance and encouragement over the years. Their example and expectations represent the highest standards I can aspire to and strive towards. None of my doctoral work would have been possible without the coauthors listed on all chapters, I am immensely grateful for all the work, inspiration, patience and graciously shared credit. I also learnt a humbling amount from Harvard faculty, fellow students, and visitors too numerous to name all.

In particular, the work presented in Chapter 1 improved thanks to comments of Rob Alessie, Joseph Altonji, Paul Bingley, Martin Browning, David Cutler, Mette Ejrnæs, John Friedman, Ed Glaeser, Thais Lærkholm Jensen, Daniel Le Maire, Søren Leth-Petersen, Daniel Prinz, and numerous seminar participants at Harvard, SFI, DGPE, and CAM. I am grateful to Torben Heien Nielsen and Tore Olsen for graciously sharing their code and tax calculations with us.

The work presented in Chapter 2 benefited greatly from comments of Nava Ashraf, David Autor, Stefano DellaVigna, Hilary Hoynes, Damon Jones, Emir Kamenica, Lawrence Katz, Henrik Kleven, Ulrike Malmendier, Monica Singhal, Timothy Taylor, and numerous seminar participants. I thank Jenny Henzen and the Elsevier staff for supporting the project and providing data and Liz Anderson for implementing the experiment. Greg Bruich, Jessica Laird, Keli Liu, Alex Olssen, and Heather Sarsons provided outstanding research assistance.

For the work presented in Chapter 3, I thank Parag Pathak and Georg Weizsäcker for participating in the earlier stages of the project, especially for Georg Weizsäcker's input in designing and running the experiment, and also the research assistants who helped us at various stages of the project: Aðalheiður Ósk Guðlaugsdóttir, Hjörleifur Pálsson, Thorhildur Ólafsdóttir, Kristófer Gunnlaugsson, Kristín Eiríksdóttir, Kári S. Friðriksson, Katrín Gunnarsdóttir, Valur Þrainnsson, Jónas Örn Helgason, and Dagný Ósk Ragnarsdóttir.

At the end of the day, usually a long day, friends kept me sane, and my family has supported me over the years of my education. Yet only my wife, Orsolya Mednyánszky,

really knows all that is behind this dissertation. Thanks for joining me on this adventure—  
there is more to come.

To my family

# Introduction

Would the value of unemployment insurance fall if more people had a buffer stock of liquid savings? Using quasi-experimental evidence from the unexpected introduction of home equity loans in Denmark, where public unemployment insurance is voluntary, Chapter 1 documents that liquidity and insurance are substitutes. A Danish reform provided less levered homeowners with more liquidity. Using a ten-year-long panel dataset drawn from administrative registries, we find that people who obtained access to extra liquidity were less likely to sign up for unemployment insurance. The effect is concentrated among those for whom insurance has negative expected value. In this group, extra liquidity from housing equity worth one year's income decreases insurance up-take by as much as a 0.3 percentage point fall in the risk of unemployment. Placebo tests for earlier years show no differential trends by leverage before the natural experiment. This implies that the liquidity of financial assets influences unemployment insurance uptake in the absence of public provision of insurance.

In Chapter 2, we evaluate policies to increase prosocial behavior using a field experiment with 1,500 referees at the *Journal of Public Economics*. We randomly assign referees to four groups: a control group with a six week deadline to submit a referee report, a group with a four week deadline, a cash incentive group rewarded with \$100 for meeting the four week deadline, and a social incentive group in which referees were told that their turnaround times would be publicly posted. We obtain four sets of results. First, shorter deadlines reduce the time referees take to submit reports substantially. Second, cash incentives significantly improve speed, especially in the week before the deadline. Cash payments do

not crowd out intrinsic motivation: after the cash treatment ends, referees who received cash incentives are no slower than those in the four-week deadline group. Third, social incentives have smaller but significant effects on review times and are especially effective among tenured professors, who are less sensitive to deadlines and cash incentives. Fourth, all the treatments have little or no effect on agreement rates, quality of reports, or review times at other journals. We conclude that small changes in journals' policies could substantially expedite peer review at little cost. More generally, price incentives, nudges, and social pressure are effective and complementary methods of increasing prosocial behavior.

Chapter 3 examines the open empirical question of whether people, when choosing among rewards received at different points of time, are influenced by their expected income levels at those times. Moreover, we seek to measure time preferences after compensating for possible income effects. Besides eliciting subjects' preference between standard delayed rewards, the experimental design also elicited their preferences over delayed rewards that are received only if the subject's income remains approximately constant. These preferences along with elicited subjective probabilities of satisfying the condition make the correction possible. We conducted the experiments in Iceland, where prompt availability of income tax returns enabled us to condition delayed rewards on income realizations. We find that background income affects preferences over unconditional delayed rewards. While most people exhibited present bias when comparing unconditional delayed rewards, subjects with stable income did not. The results are similar for the entire sample once we correct subjects' discount functions for income effects. This suggests that income expectations have an effect on choices between future rewards, and that this may account for some of the present-bias observed in experiments.

## Chapter 1

# Does Liquidity Substitute for Unemployment Insurance? Evidence from the Introduction of Home Equity Loans in Denmark<sup>1</sup>

### 1.1 Introduction

To what extent is liquidity a substitute for insurance? The answer to this question has important implications for the design of optimal social insurance policies. If relaxing liquidity constraints enables people to better smooth marginal utility and address their specific needs, credit could be a partial substitute for government insurance, attenuating the typical distortions and fiscal externalities of traditional tax-and-benefit schemes.

Unemployment insurance is a classic example of social insurance, so much so that in most countries public unemployment insurance schemes are standardized and mandatory. While Denmark has a public unemployment insurance scheme, individual participation

---

<sup>1</sup>Co-authored with Kristoffer Markwardt and Alessandro Martinello

is voluntary. Because the supply of unemployment insurance is fixed, as this scheme is standardized and publicly regulated, we are able to study the demand for insurance by looking at changes in subscription rates. In this paper, we bring quasi-experimental evidence on the insurance choices of 113,000 homeowners in their late twenties and thirties, and test whether those who were allowed to borrow against equity in their homes bought differentially less unemployment insurance afterwards.

Prior to an unanticipated 1992 mortgage reform, borrowing against home equity from mortgage banks for consumption purposes was illegal in Denmark. By introducing home equity loans, the reform unexpectedly provided some homeowners with extra liquidity, without any differential change in their wealth. This motivates what is essentially a difference-in-differences research design. Combining this approach with a ten-year panel dataset drawn from administrative registers, we find that liquidity does substitute for unemployment insurance. More specifically, homeowners who had a substantial amount of home equity at the time the reform was enacted subscribed relatively less to unemployment insurance funds after the reform, compared to homeowners with little or no accessible home equity. We find that an increase in accessible liquidity worth one year of income caused about one Dane in two hundred to forgo unemployment insurance. The effects are concentrated among the group whose insurance is not actuarially fair; a year's income's worth of extra liquidity reduces their insurance up-take by 0.94 percentage points. This is equivalent to the effect of a 0.3 percentage point, or 15%, decrease in their estimated risk of unemployment. We show that groups with higher unemployment risk show little or no response. Our placebo tests validate our design, as home equity is not correlated with differential trends in insurance prior to the reform.

The Danish institutional features put the magnitude of the effect into context. First, we document substantial persistence in unemployment insurance membership and high baseline insurance up-take, a finding confirmed by Parsons *et al.* (2003). Only 13% of the individuals in our sample change insurance status over the ten years covered by our data. This type of persistence can stem from the psychological costs of changing insurance status,

from the relative generosity of the scheme, or from social norms of solidarity. Second, home equity loans in Denmark carried large transaction costs compared to a HELOC (home equity line of credit) in countries such as the United States: The process required interviews at the issuing bank, and even after a line of credit had been established, the borrower could not freely draw upon home equity with a credit card, but instead had to apply for each additional loan. Hence, the 1992 reform likely had a smaller impact than it would have had under the more permissive loan policies in place in other countries.

We contribute to the literature on social insurance by estimating the extent of the substitution between formal insurance and a buffer stock of savings, a crucial quantity for the design of optimal unemployment insurance schemes.<sup>2</sup> From empirically verified models of lifetime consumption, we know that agents should and do accumulate liquidity as a means to smooth consumption. Precautionary savings respond to income risk (Carroll, 1997, 2009), liquidity constraints (Alessie *et al.*, 1997; Deaton, 1991), and commitment constraints (Chetty and Szeidl, 2007). Engen and Gruber (2001) show that more generous public insurance schemes crowd out private savings; reducing the unemployment insurance replacement rate by 50% in the U.S. would increase gross financial asset holdings by 14%. This is true even though savings do not help transfer wealth from the more fortunate to those with more or longer unemployment spells. Yet by cushioning the blow in the bad state, buffer stocks can limit the need for state contingent claims, and are thus often called self-insurance.<sup>3</sup> If people are unable to smooth consumption and are forced to cut back spending in unemployment (Gruber, 1997) then insurance schemes with high replacement rates are optimal (Hansen and Imrohorouglu, 1992; Crossley and Low, 2011; Lentz, 2009). With unconstrained borrowing however, much smaller benefits are optimal as residual insurance against the duration of

---

<sup>2</sup>Davidoff (2010) shows that home equity does limit the demand for long-term care insurance. In his framework, the house is sold before moving to a nursing home, which does yield more resources to pay for care, yet does not show that a liquid buffer and anticipated intertemporal smoothing would limit the need for insurance.

<sup>3</sup>If all unemployment risk were within person, i.e. only the spells' timing were unpredictable but not the overall exposure over a lifetime, free intertemporal smoothing would eliminate the need for any formal insurance. Even if there is a cross-sectional component of risk across individuals, facilitating the reallocation of one's own resources over time decreases the importance of insuring an unlucky career against luckier ones.

the unemployment spell (Shimer and Werning, 2008).

Empirical work on distinguishing between the liquidity and the moral hazard effects of unemployment benefits also document that the former overwhelms the latter (Chetty, 2008). Our result suggests that the accumulation of liquidity is a real alternative for some households; even preferable to paying the premium on an unemployment insurance contract.<sup>4</sup> While mandatory savings accounts have been speculated to be a more efficient alternative to conventional insurance (Feldstein and Altman, 2007), and are currently in place for instance in Chile or in Singapore, we have little quasi-experimental evidence on their effects (Chetty and Finkelstein, 2013).

Evidence from the standardized and subsidized Danish market, however, is not without limitations: In the absence of endogenous pricing, let alone contract design as discussed by Hendren (2013), the Danish experience does not prove that an unsubsidized unemployment insurance market is viable with tight liquidity constraints, nor how much such a market would unravel if the constraints were relaxed. This paper documents one channel for partial unraveling, which is likely to exacerbate adverse selection in an open market.<sup>5</sup>

We do not observe home equity after 1992 and thus cannot perform an analysis of ex post responses to shocks to verify that home equity serves as a liquid asset. However, there is evidence that people, given the opportunity, draw upon their home equity to finance consumption. Hurst and Stafford (2004) find that households borrow against their home equity in periods of unemployment. Moreover, using the same reform, a liquidity shock has been found to increase consumption over time (Leth-Petersen, 2010), and even encourage entrepreneurship (Jensen *et al.*, 2014). Finally, Chetty and Szeidl (2010) show that higher

---

<sup>4</sup>This does not contradict the fact that these people did not accumulate liquid buffer stocks in the old regime, when only specific instruments were liquid.

<sup>5</sup>Unemployment insurance markets could unravel if only inherently risky groups insure themselves (adverse selection), or once insured, people are less careful to keep their jobs and get back to work quickly (moral hazard). This paper suggests that people can also select out of insurance if they foresee their ability to take precautions, which would exacerbate adverse selection. Einav *et al.* (2013) document similar selection on an anticipated behavioral response, that more price sensitive patients select into plans with lower copays, which they call selection on moral hazard in health insurance. Our data limits us from gathering direct evidence on whether people with overpriced insurance invested more in home equity after the reform to grow a buffer.

property values, holding wealth equal, lead to increased tolerance for risk; an increase in home equity increases the probability of investing in the stock market. These empirical findings all support our interpretation of home equity as an imperfectly liquid asset, once home equity loans are allowed.

The paper proceeds as follows. Section 1.2 introduces the 1992 credit market reform and the Danish unemployment insurance system in more detail and outlines our empirical strategy. Section 1.3 describes our data, motivates our sample selection, and provides summary statistics. Section 1.4 presents our results and discusses their robustness. Section 1.5 concludes.

## **1.2 Institutional Details and Empirical Strategy**

This paper identifies the effect of liquidity on the demand for public unemployment insurance by exploiting a large, sudden, and unexpected policy variation and the features of the Danish unemployment insurance system. The mortgage reform, which was approved by the Danish parliament in May 1992, unexpectedly endowed some homeowners with extra liquidity. The voluntary nature of Danish unemployment insurance enables us to study its demand. This section of the paper describes how we exploit these two features of the Danish system to identify the effect of liquidity on the demand for unemployment insurance.

### **1.2.1 The Danish Unemployment Insurance System**

The Danish unemployment insurance system builds upon several unemployment insurance funds (42 in 1992), which are private associations of workers with the purpose of providing economic support to their members during unemployment. However, as funds are complemented by the state—the system is self-supporting only if the unemployment rate is around 3%—strict regulation at the national level demands that each fund offers a uniform insurance product, independent of the occupation and industry of its members. While unemployment benefits are thus identical nationwide, Danish workers are free to choose

whether to subscribe to an unemployment insurance fund or bear the risk of unemployment themselves.

These funds are the main, but not the only, source of income contingent on losing a job. While a publicly funded welfare program exists, supplemental security income eligibility requirements are very strict. In principle, applicants cannot own any assets, or be able to sustain themselves in any other way (for example through another earner in the household), in order to be considered for welfare benefits, which are also lower than those received from unemployment insurance funds.<sup>6</sup> Because no major changes in the supplemental security income system occur during the period of interest, we ignore welfare benefits in our analysis. Benefit amounts are detailed in Appendix Table A.4.

Wage earners and the self-employed have access to different unemployment insurance schemes in terms of eligibility rules and requirements once unemployed. For wage earners (about 90% of the Danish workforce), to whom we restrict our analysis, eligibility for receiving unemployment insurance benefits requires uninterrupted membership in an unemployment insurance fund in the 12 months preceding unemployment and at least 26 weeks of paid work over the last three years. The funds do not screen applicants for membership. Special rules apply to recent graduates, who are immediately eligible for unemployment insurance benefits if they sign up within one month from graduation.<sup>7</sup>

To retain benefits, the unemployed must comply with a set of rules, specified in ministerial guidelines on active labor market policies. These guidelines require recipients to make their resumes publicly available, apply for at least a given number of jobs per month, and participate in courses and other activities assigned by their caseworker on the basis of individual abilities and potential. Under these criteria, the daily benefits can amount to up to 90% of their daily gross income averaged over the preceding 12 weeks. However, the

---

<sup>6</sup>Probably the largest variation in the value of the unemployment insurance contract comes from changes in marital status, which changes supplemental security income eligibility. We do not model this explicitly, but, as a robustness check, we investigate separately households with constant marital status around the reform in the appendix.

<sup>7</sup>Students receive reduced benefits the first year, which corresponds to approximately 80% of the standard benefit level.

benefits are capped at a relatively low level. In 1992, the cap on benefits corresponded to a gross monthly salary of approximately \$2,000, and thus affected 95% of full-time insured workers.<sup>8</sup>

The unemployment insurance contract is a bargain for most Danes, unless they have very low (subjective) unemployment risk or high hassle costs of the contract. The yearly statutory membership fee was the equivalent of eight times the maximum daily benefits over this period, e.g. eight times 417 DKK in 1992 for full-time workers. Hence, in absence of additional administrative fees and taxes, the insurance would be actuarially fair for workers facing duration-weighted unemployment risk of about 2.6% per year.<sup>9</sup> Benefits count towards taxable income and every person  $i$  in year  $t$  can calculate the expected value of future benefits using their future retention rate,  $(1 - MTR_{i,t+1})$ .<sup>10</sup> Membership fees are tax deductible, but only from a special notion of taxable income, where the top two tax brackets do not apply. Thus the relevant retention rate for the fees is  $(1 - MTR_{i,t+1}^{bottom})$ . The expected net benefit of membership after taxes is a multiple of daily benefits ( $DB$ ):

$$NB_{i,t} = \left[ (1 - MTR_{i,t+1}) \cdot YED_{t+1} \cdot UR_{i,t+1} - (1 - MTR_{i,t+1}^{bottom}) \cdot 8 \right] \cdot DB_{t+1}, \quad (1.1)$$

where  $YED$  is the full-time, full-year (FTFY) equivalent number of days (312 in 1992) and  $UR$  is unemployment risk as a fraction of the year spent unemployed. We plot our estimates about expected net benefits for the following year by levels of (estimated) unemployment risk in our estimation sample, using 1987 as an example, in Figure 1.1. Note that the subsidized insurance scheme is a lottery with positive net expected value for many, though our calculation probably overestimates net benefits as we can calculate risk (FTFY benefit take-up) only on the insured, who are subject to adverse selection, moral hazard, and selection on moral hazard.

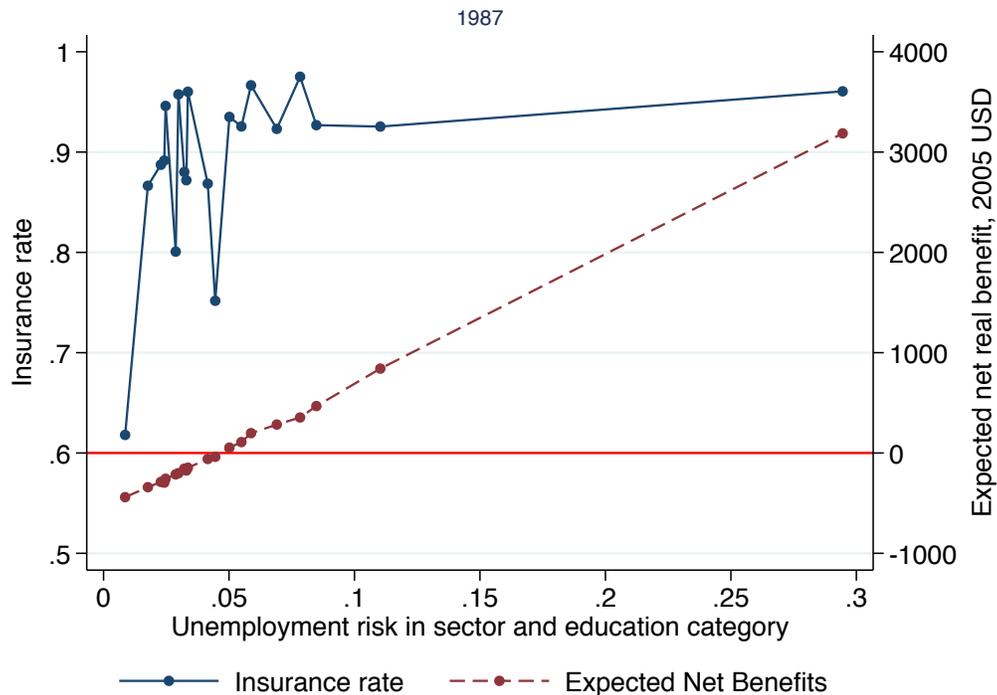
---

<sup>8</sup>Apart from the documentation of rules, we report all monetary amounts in US dollars, using the 1991 exchange rate of 5.91, while also correcting for domestic inflation, using 2005 prices for more familiar magnitudes.

<sup>9</sup>E.g. a 2% Bernoulli risk of spending half a year on benefits corresponds to 1% risk for the calculation of expected benefits.

<sup>10</sup>This calculation ignores the fact that spells long enough for people close to thresholds of tax brackets could knock them down into a lower tax bracket.

**Figure 1.1:** *Expected Net Benefits of Unemployment Insurance in the Estimation Sample*



NOTE.— The figure presents average insurance purchase (membership in November) and our calculated expected net benefits (dashed line) against average full-time full-year equivalent unemployment risk in twenty equal-sized bins, for the 1987 insurance decision. Both series are plotted as means in 20 equal sized bins by risk, connected for illustration. The net benefit is expressed in 2005 US dollars (2005 DKK values using the domestic CPI, converted to USD using the 1991 exchange rate of 5.91). Unemployment fund membership is measured in November 1987 but coverage applies to 1988. The marginal (bottom) tax rate used for net benefits come from each taxpayer’s actual MTR in 1988, according to our calculation based on observed incomes and determinants of the tax schedule. Unemployment risk here is the average FTFY equivalent time spent on benefits in 1988 for others in the estimation sample who are full-time insured in the same industry and broad education category in November 1987. This leave-out mean unemployment risk predicts realized unemployment with an  $R^2$  of 0.59 over the 1987-1995 period. In 1988, FTFY unemployment corresponded to 312 days of the daily maximum benefits, and membership fees to 8 days worth of benefits. This calculation does not use the 90% replacement rate for those who do not hit the benefit cap. See equation (1.1) for the specific formula used.

Appendix Table A.5 collects the relevant parameters of unemployment insurance over our time period. The fairly high benefit level combined with low after-tax insurance fees makes unemployment insurance attractive for many. Meanwhile social norms and inertia together with the historically tight bond between unemployment insurance funds and labor unions imply high insurance up-take in Denmark. The characteristics of the unemployment insurance market in Denmark allow us to study the unemployment fund membership of Danes as a proxy for their effective risk tolerance. As the supply of insurance is fixed and publicly regulated, the market outcome is determined only by the demand for insurance.

Unemployment insurance fund membership is, however, not entirely driven by demand for insurance against job loss, but also by eligibility for an early retirement scheme (*efterløn*), which allows members to retire at age 60 rather than at age 67 (the official retirement age during the period of interest). Many Danes take advantage of this possibility to retire early. Approximately 50% of the population received *efterløn* at the age of 64 between 2007 and 2011, which is halfway between the earliest eligible age of 60 and 67, when public pensions become available.

Until 1992, eligibility for early retirement benefits required membership of an unemployment insurance fund for the last ten years before retirement; then this requirement increased to twenty years. People between the ages of 40 and 50, who were not already unemployment insurance fund members, were given the option to join no later than March 1992 to acquire eligibility at age 60. Many people in their forties committed to the scheme in 1992, and hardly made a choice about insurance ever after.<sup>11</sup> Those who did not join that year constitute a self-selected group, some of whom still joined later to enjoy early retirement at an age older than 60. We restrict our analysis to younger cohorts, unaffected by this change. A detailed account of how older cohorts are affected by early retirement reform is deferred to Appendix Table A.2.

---

<sup>11</sup>The reform might have changed ex post behavior for those who found themselves insured for this unrelated reason. Ejrnaes and Hochguertel (2011) attribute different self-employment patterns at different ages to this shift of ten cohorts into the funds in 1992, also insuring some business risk of sole proprietors.

### 1.2.2 The Danish Mortgage System and the 1992 Reform

Most real estate purchases in Denmark are financed via mortgage credit institutions, which offer loans with the property as collateral.<sup>12</sup> The legal cap on loan to value (LTV) is 80%, the homeowner must provide a 20% downpayment. Mortgage credit institutions issue callable bonds to fund pools of loans, and the securitized loans are thus low-risk and highly liquid. Real estate loans are cheaper than personal loans established with commercial banks after a credit review, especially after losing a job and without collateral. Denmark has no national credit bureau, and few workers could expect to have a line of credit open, mainly in the form of a credit card, after being laid off.

In 1992, Folketinget, the Danish parliament, voted in favor of a mortgage reform, shortly after a brief discussion in the spring. Before May 21, 1992, Danes could get a securitized mortgage only for real estate investments (purchase or remodeling). Thus home equity used to be a highly illiquid asset, which could be turned into cash only through a sale or perhaps a costly and uncertain loan. The reform changed mortgage regulation in three ways: maximum maturity, remortgaging, and the use of the loan. The last is the crucial element for the purpose of this paper; allowing mortgage loans to finance purposes other than real estate investments effectively let Danes to use up to 80% of their real estate wealth as collateral for consumption loans established through mortgage credit institutions.<sup>13</sup>

The reform was unanticipated; Leth-Petersen (2010) has documented that not even the major finance and economics newspaper in Denmark covered the reform until the month it was enacted. This unanticipated access to credit of particular homeowners allows us to isolate the causal effect of an increase in liquidity on the demand for formal insurance, holding wealth fixed: Households did not hold more or less home equity at the time of the reform because they anticipated its use as a liquid buffer stock.<sup>14</sup>

---

<sup>12</sup>On general features of the Danish mortgage market, including their implementation of the European covered bond system, see Campbell (2013).

<sup>13</sup>The limit was initially set at 60% but was quickly raised to 80% by December 1992.

<sup>14</sup>For the thought experiment behind our causal reasoning, wealth should not change along with liquidity. In our quasi-experiment, most of the sample became wealthier in the years following the reform, but not

After 1992, turning home equity cheaply into cash on hand still required a new mortgage contract with non-trivial transaction costs, yet homeowners were no longer forced to sell a house and move just to tap into this asset. However, the liquidity of homeowners with a large established mortgage did not change, because they could not mortgage up any more than they already had. They can thus serve as a control group, making our identification strategy straightforward. We compare homeowners endowed with home equity just before the reform to homeowners who mortgaged to the limit, and estimate how their insurance choices evolved differently over time. Our specification is therefore similar to a difference-in-differences design.

The 1992 reform has two key elements that make the Danish case uniquely valuable to identifying liquidity effects. First, the reform was unanticipated. Because people could not know that the reform was to be implemented, they could not have adjusted their housing, insurance and liquidity accordingly, and no voluntary selection into treatment could have taken place. As we have access to data on homeowners from 1987, five years before the reform took place, we can show that the trends in insurance up-take were identical for the treatment and the control groups (high and low equity owners, respectively) up to when the reform was implemented.

Second, the reform only changed the costs of tapping into home equity, but did not affect individual wealth differently for those with more or less home equity. Therefore, we are able to identify the liquidity effect on insurance demand, independently of wealth. This unique feature of our identification strategy distinguishes this paper from those studying behavioral effects of changes in wealth (Shapiro and Slemrod, 2003; Chetty and Szeidl, 2010; Andersen and Nielsen, 2011), and is more directly comparable to studies of direct liquidity shocks (Gross and Souleles, 2002).

However, the reform also changed mortgage regulations in two other ways, namely by introducing the right to cash-out refinancing and by expanding the maximum maturity

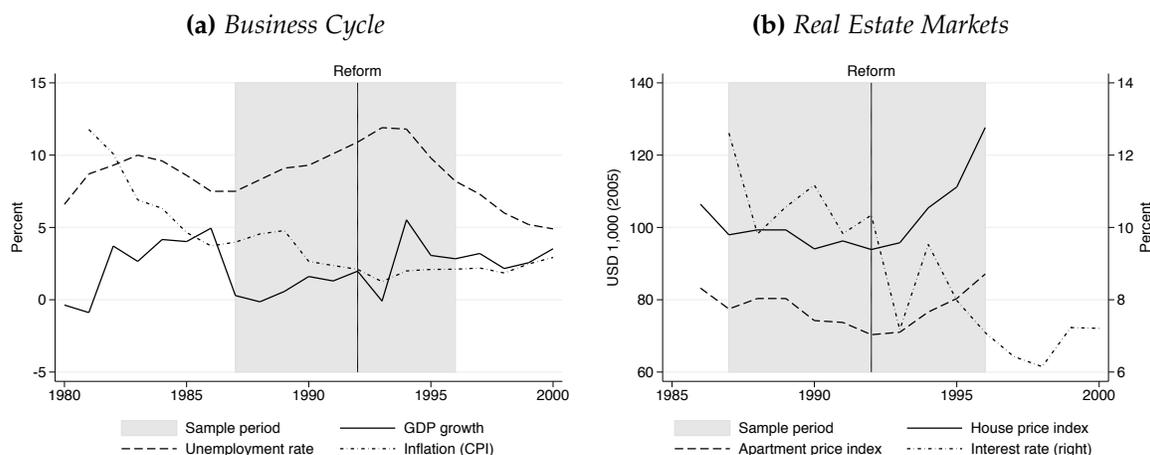
---

differentially for those with more home equity, thus a higher dose of treatment. As our summary statistics in Table 1.1 show, home equity is not correlated with total housing wealth in our sample, so house price rises after 1991, be they secular or an effect of the mortgage reform, made everyone equally wealthier.

of real estate loans from 20 to 30 years. Remortgaging gives the debtor the possibility to lower the cost of his debt when market interest rates fall. A borrower is entitled to redeem a mortgage bond at par at any time prior to maturity by prepayment, and thus to exploit interest rate changes to reduce the costs of funding. Because interest rates were falling on average during our sample period—shown in the right pane of Figure 1.2—this opportunity was particularly valuable for holders of large mortgages. Though the option value to remortgage constitutes a wealth transfer to our control group, it is annuitized, as remortgaging changes monthly installments, and thus is hard to turn into cash on hand. With no equity in their homes, even the more flexible refinancing options after 1992 do not allow our control group to get cash out from refinancing.

Figure 1.2 shows that between 1987 and 1993 Denmark suffered a period of economic stagnation, with rising unemployment rates. Moreover, real estate prices changed considerably in our period of interests, both for apartments and houses, and generally increased after 1993. Because of this economic turmoil, we exploit not only the long panel structure

**Figure 1.2: Economic Environment**



NOTE.— Real estate prices reflect market transactions. Interest rates refer to annual average yields of 20-year maturity mortgage-credit bonds. 2005 US dollar values are 2005 DKK values using the domestic CPI, converted to USD using the 1991 exchange rate of 5.91.

Sources: For one-family home and apartment prices, Statistical Yearbook (*Statistisk Årbog*), 1988-1998; For other variables, Statistics Denmark ([www.statistikbanken.dk](http://www.statistikbanken.dk), NAT02 B1.\*g, AULAAR, PRIS12 and DNRENTA series, accessed on December 5, 2012).

of our dataset to control for year-specific fixed effects, but also the richness of Danish administrative registers to control for several demographic and financial characteristics.

### 1.2.3 Econometric Methods

Our general specification takes the form of the linear probability model

$$I_{i,t} = \alpha_{\text{HE}} \cdot \text{HE}_{i,1991} + \tau \cdot \text{HE}_{i,1991} \cdot \mathbf{1}[t \geq 1992] + \mathbf{X}_{i,t}^d \boldsymbol{\beta}_d + \mathbf{X}_{i,t}^f \boldsymbol{\beta}_f + \nu_t + u_{i,t} \quad (1.2)$$

where  $I_{i,t}$  indicates insurance status in year  $t$  for individual  $i$ . The linear probability model in equation (1.2) is similar to a difference-in-differences design, where the coefficient  $\tau$  identifies the change in average enrollment in an insurance fund for any given mortgageable home equity  $\text{HE}$ , as measured in December 1991, relative to secular time fixed effects  $\nu_t$  and how (1991) home equity correlates with insurance demand in the cross section,  $\alpha_{\text{HE}}$ . This model includes financial and demographic controls,  $\mathbf{X}_{i,t}^f$  and  $\mathbf{X}_{i,t}^d$ , and allows for arbitrary correlation within individuals in the residuals  $u_{i,t}$ . We normalize home equity and financial controls by permanent income.

We also estimate standard difference-in-differences models with no scaling by treatment dose, with a treatment and a control group based on the amount of home equity held in December 1991 relative to our permanent income measure. We assign to the control group those holding no mortgageable equity, and to the treatment group those who hold more than a month's income in home equity. We argue that those who owned a home at the beginning of 1992 but had too little equity to take advantage of the new rules and those who held more home equity before 1992 were experiencing common underlying trends when it came to unemployment insurance. Under this assumption, the difference in changing insurance behavior from 1992 onwards is caused only by the availability of home equity loans from mortgage credit institutions.

While this definition of the control group reflects important individual choices before 1992, the reform was unexpected, thus those choices could not be motivated by the need of extra liquidity. Our identifying assumption is that the underlying differences between the

treatment and the control group do not drive different trends in insurance purchase. We show that the trends in unemployment insurance fund memberships were identical in the two groups before the reform.

As we do not observe home equity after 1992, our definition of treatment and control groups allows for some in the control group gaining some treatment over time, as house prices rise and mortgages are paid down. As the initial difference in liquidity does not diminish, our preferred interpretation is that more liquidity causes less demand for insurance, with little to say about the nonlinear effects of having any or little liquidity. The parallel trends of uniformly improving liquidity is important for the interpretation of either specification. If the home equity gains after 1992 are more valuable for our control group than our treatment group, this attenuates our estimates of the long-run effect of liquidity.

Using initial home equity as an instrument for equity in later years would essentially correct our estimates for measurement error and imperfect compliance in slack liquidity constraints later (maybe partly in response to the reform). We interpret the reduced-form estimates as a lower bound for the substitution between private precautionary savings and formal unemployment insurance.

Unemployment insurance choices are characterized by strong inertia. Only about 13% of the individuals in our sample change insurance status at least once during the ten years of our analysis. This persistence in insurance choice comes not only from the eligibility criteria for unemployment insurance benefits, which encourage continued enrollment, but also from the historically strong connection between unemployment insurance funds and labor union memberships, social norms of solidarity, and psychological costs of changing insurance status.

We model this inertia in two alternative ways. First, we include individual fixed effects in some specifications, thereby identifying the parameters using only variation among those who change their insurance status. Second, we estimate a lagged dependent variable model, which has a particularly meaningful interpretation in a random utility framework. In this setup, which in its simplest form corresponds to a standard logit model, an agent subscribes

to an unemployment insurance fund if the utility of being insured is larger than the utility of being uninsured. Essentially, we assume a random utility model

$$u_{i,I,t} = v_{i,I,t} + \varepsilon_{i,I,t}, \quad I \in \{0, 1\}, \quad (1.3)$$

where we model the predictable utility of insurance status  $I$  for individual  $i$  at time  $t$ ,  $v_{i,I,t}$ , as a linear combination of observables, while  $\varepsilon_{i,I,t}$  is unobserved and follows a logistic distribution. The individual chooses to be a member of an insurance fund if  $u_{i,1,t} > u_{i,0,t}$ . Thus,

$$\Pr_i(I_t) = \frac{1}{1 + \exp(v_{i,1-t} - v_{i,I,t})}. \quad (1.4)$$

To model inertia, we assume that the agent pays a one-time utility cost  $c_1$  for subscribing to an unemployment insurance fund and a parallel cost  $c_2$  to unsubscribe from the fund. These costs can reflect administrative fees, opportunity costs, or simply the psychological effort of gathering information on how to change one's fund membership and submit the necessary paperwork. The non-random part of individual utility will then be state dependent and (without loss of generality absorbing the relative effect of observables  $\mathbf{X}_{i,t}$  in the status of insuring) equal to

$$v_{i,0,t} = -c_2 I_{t-1}, \quad v_{i,1,t} = \alpha + \mathbf{X}_{i,t} \boldsymbol{\beta} - c_1 (1 - I_{t-1}) \quad (1.5)$$

and the probability of  $I = 1$  is

$$\Pr_i(I_t = 1) = \frac{1}{1 + \exp[-(\mathbf{X}_{i,t} \boldsymbol{\beta} + \alpha - c_1 + I_{t-1}(c_1 + c_2))]} \quad (1.6)$$

In this model,  $\alpha$ ,  $c_1$ , and  $c_2$  are not separately identified. To see this, suppose  $(\alpha^*, c_1^*, c_2^*)$  maximize the likelihood. Then the set  $(\alpha', c_1', c_2') = (\alpha^* + k, c_1^* + k, c_2^* - k)$  yields the same likelihood for any  $k \in \mathbb{R}$ . In a standard lagged dependent variable logit model, the coefficient of  $I_{t-1}$  will then identify the sum of the two costs  $c_0 = c_1 + c_2$ . As follows from equation (1.6), the larger the switching costs, the larger the difference between previous members' and non-members' insurance up-take, as more people renew a membership or still do not join. Therefore, given the amount of inertia in the data, we expect the coefficient of  $I_{t-1}$  to

be positive and significant.

### 1.3 Data and Summary Statistics

We draw data from Danish administrative records, which are linked at the individual level. They hold detailed information on individual background characteristics, family composition, labor market attachment, insurance status, income, and wealth. The registers all provide longitudinal information on the entire Danish population, mainly at an annual frequency. The tax authority records provide detailed data on total taxable income and transfers as well as taxable wealth from 1987 to 1996 because of a wealth tax that was in effect over this period. The wealth tax implied third-party reporting of both income and wealth holdings by banks and other financial intermediaries to the tax authorities. Thus, the data we use for our empirical analysis span those ten years.

The mortgage reform in 1992 allowed homeowners to finance non-housing consumption up to 80% of the property value from mortgage credit institutions. Therefore, we use the last observation before the reform to calculate the unexpected liquidity shock by taking 80% of housing wealth and subtracting mortgage debt. The tax registers report the publicly assessed housing value by December 31 each year, which takes into account only objective and easily observable characteristics. However, a home equity loan is granted on the basis of the market price of the property. To better reflect the market fluctuations in real estate prices, we follow Leth-Petersen (2010) and use aggregate data on market transactions to adjust the observed property values by the ratio between market prices and public evaluations for each year and municipality. Each mortgage is recorded in our data as a snapshot of the market value of its callable mortgage bonds, taken on December 31.<sup>15</sup> This is the value that counts towards LTV limits on new loans.

We normalize the liquidity shock and financial controls by a proxy for permanent income, the 22-year average of real earnings from 1987 to 2008. We use as many earnings

---

<sup>15</sup>Mortgage debt is reported separately only until 1992. This limits the scope for supplementary investigations of post-reform behavior, e.g. whether they take out a home equity loan in case of unemployment.

observations as possible for this calculation to reduce the risk of comparing individuals on different parts of their life-cycle earnings trajectory.<sup>16</sup> The normalized liquidity shock we use throughout the paper is thus given by

$$L_{1991} = \frac{0.8 \times H_{1991} - M_{1991}}{Y^P} \quad (1.7)$$

where  $H_{1991}$  and  $M_{1991}$  denote housing and mortgage values as of December 31, 1991, respectively, and  $Y^P$  is annual permanent income.

We measure insurance against adverse labor market outcomes by membership in an unemployment insurance fund. The administrative records provide annual information on individual membership status by December 31 reported directly by the unemployment insurance funds.

The mortgage reform coincided with the sudden increase in the incentive to join an unemployment insurance fund for early-retirement purposes, which invalidates our identification strategy for the affected birth cohorts. Therefore, we restrict the estimation sample to individuals, who were between ages 25 and 39 throughout the period of interest, old enough to exhibit non-trivial housing and insurance choices and for whom early retirement motives did not affect the demand for unemployment insurance membership. Thus, our initial sample consists of homeowners in 1992 from the cohorts born between 1957 and 1962. As those over 35 might have joined early to count towards the retirement criterion in case they missed some (at most five) years before 60, in Appendix Table A.3 we show that our results are robust to using only the youngest of cohorts.

The housing and mortgage information used to calculate the liquidity shock reflects the values by December 31, 1991, five months before the mortgage reform in late May, 1992. To ensure that our estimates are not confounded by variation from homeowners who choose to move, and thus refinance, before the reform took effect, we exclude individuals who moved

---

<sup>16</sup>We assume away moral hazard in the earnings process. Post-reform wage earnings potentially are affected by reoptimization due to the changed portfolio composition caused by the reform. However, we do not regard this as a substantial issue compared to the improvement in the approximation to permanent income that these extra years provide. Restricting the measure to pre-reform earnings does not change our results qualitatively.

within the first five months of 1992, according to residence records.

Our financial variables directly reflect individual tax forms from third-party reports. Irregularities may or may not have been corrected; as most Danes have too little wealth to be taxed, neither has the tax agency any incentive to correct underestimates, nor the taxpayer to correct overestimates. Because the wealth data comes from snapshots as of December 31, imprecisions in the timing of the reports can affect what the researcher observes in the data. Such transitory irregularities are unavoidable in public administrative records and introduce noise in the financial variables. However, we exclude persistent outliers, for example, the very rich. An individual is excluded from the sample if, for at least one of the financial variables (housing wealth, mortgage debt, assets, liabilities, disposable or permanent income), his average value over the entire sample period is in the top 1% of the distribution. We further condition on participation in the labor force and being a wage earner (the self-employed have a different unemployment insurance scheme). We also exclude records with incomplete information on labor market attachment such as industry code or experience.

Buying and selling real estate involves several transactions that are potentially executed and registered at different points in time. Because housing and mortgage values are snapshots on December 31, 1991, a real estate transaction close to that date potentially implies that these values refer to different pieces of property. While such patterns are obvious for some observations, we are unable to systematically identify these errors because both values fluctuate from year to year, and people may buy either a new home or another home. We exclude individuals for whom we calculate a liquidity shock in the top or bottom 1% of the distribution.

Finally, we restrict our analysis to individuals who are observed in all years between 1987 and 1996. We keep a fully balanced sample to avoid changes in sample composition due to attrition by migration or death. As we do not model how Danes plan their insurance membership when they go on parental leave, back to school, abroad, or into self-employment, the fully balanced sample also implies that we document the substitution between liquidity

and insurance in the self-selected subgroup who remain employed (or unemployed) from 1987 to 1996. This might affect the external validity of our findings, e.g. the population treatment effect including post-1992 entrepreneurs' need of insurance might be higher. The internal validity of our difference-in-differences design is not under threat in the fully balanced sample.

Our final sample consists of 113,344 individuals, detailed in Appendix Table A.1. We compare summary statistics of this sample to the entire population of the same cohorts in Table 1.1. Columns 2-5 divide the selected sample of homeowners in 1991 into quartiles of the liquidity shock induced by the reform, while the last column reports values for the entirety of the six cohorts.

The table reports the liquidity shock and its subcomponents (housing wealth, mortgage debt, and permanent income) as well as the evolution of the insurance up-take, move-in date, and socioeconomic variables. In addition, labor market attachment is characterized by disposable income (total current-year income net of taxes), accumulated labor market experience over the past five years, and individual unemployment risk. The unemployment risk is given by the following year's industry- and occupation-specific unemployment rate.<sup>17</sup> Financial variables include liquid assets net of stock holdings, which are very noisily recorded in the registers, and total debt net of mortgage debt.

The table shows that the reform changed the liquidity of less than half the homeowners in the sample. The average amount of extra liquidity gained by homeowners in the top quartile is two thirds of annual permanent income, whereas homeowners in the bottom quartile were far from being able to use their real estate as collateral for personal loans from mortgage credit institutions. The time trends in insurance up-take show that those who are affected more by the reform generally bought more unemployment insurance in the first place.

The median housing values do not vary much across the quartiles of the liquidity shock;

---

<sup>17</sup>As we do not observe occupational level after 1995, we cannot compute this measure for 1996, and we therefore exclude 1996 from our conditional analysis.

**Table 1.1:** Summary Statistics for the Estimation Sample in 1991 by Liquidity Shock Quartiles, Compared to the Danish Population of the Same Birth Cohorts

	Q1	Q2	Q3	Q4	Population
Liquidity shock (see text)	-1.07	-.37	-.02	.69	
Insurance rate 1989 (%)	85.7	87.4	89.8	89.7	74.5
Insurance rate 1991 (%)	86.5	88.2	90.4	90.3	75.5
Insurance rate 1993 (%)	89.8	91	92.6	92.7	79.8
Moved to 1991 housing (year)	1987.3	1987.1	1986.5	1984.9	
Financial variables (2005 USD)					
Housing wealth	\$63,688	\$58,444	\$58,217	\$64,265	
Mortgage debt	\$84,479	\$59,544	\$47,441	\$33,654	
Permanent income	\$34,360	\$34,738	\$34,097	\$32,831	\$28,259
Disposable income	\$33,883	\$32,868	\$31,943	\$30,936	\$27,786
Liquid assets	\$2,850	\$2,915	\$2,971	\$3,275	\$1,847
Debts	\$14,819	\$15,893	\$16,041	\$16,252	\$9,206
Labor market measures					
Employment rate (%)	97.7	97.8	97.3	97.2	75.4
Experience, 1987-91 (years)	4.6	4.6	4.6	4.5	3.4
Unemployment risk (%)	8.4	8.5	8.8	9.1	8.9
Industry, fewest	Fi	Mi	Fi	Fi	Mi
Industry, most	Me	Me	Me	Me	So
Demographic information					
Age	31.6	31.5	31.6	31.9	31.5
College graduates (%)	22.7	22.1	20.5	16.5	24
Married or cohabiting (%)	64	59.8	59.8	60.5	47.6
Number of kids	1.3	1.2	1.3	1.3	1.1
Female (%)	44	37.4	35.6	42	49.2
Observations	28,336	28,336	28,336	28,336	439,016

NOTE.—Industry codes (NACE rev. 1): Fi: Fishing (B); Mi: Mining (C); Me: Metal industry (D); So: Other community, social and personal service activities (O). The financial variables are reported as medians in 2005 US dollars (2005 DKK values using the domestic CPI, converted to USD using the 1991 exchange rate of 5.91). Because renters and people living with parents are included in the population (column 5), we do not report housing and mortgage values for this group. The estimation sample comprises homeowners between 25 and 35 in 1991 – for more detailed information about sample selection, see Section 1.3.

they are only slightly higher in the top and bottom quartiles than in the middle two. This implies that the variation in the liquidity shock comes from differences in the mortgage values, which indeed differ substantially: The median mortgage value decreases by more than \$10,000 from each quartile to the next; the largest decrease being \$25,000 from the bottom quartile to the next. Many borrowers are under water in 1992 after recent declines in interest rates.

Much of the variation in mortgages is a result of people settling down at different points in time. This is consistent with what we observe about the time spent in the house they live in in 1991: The longer one had already lived there, the smaller was the still outstanding debt. All other variables are fairly stable across quartiles of the liquidity shock. This supports the intuition that the variation in home equity holdings is primarily caused by timing of purchase rather than selection on observed characteristics.

The sample of homeowners differs from the general population in their unemployment insurance and employment rates. Both of these differences could, however, potentially be attributed to those out of the labor force. Students, whom we exclude, have no incentive to buy unemployment insurance before graduation, while they are included in the calculation of the employment rate. Danes out of the labor force or renters imply the differences in other variables such as income, assets, and debt. Furthermore, as people who stay longer in the educational system tend to settle down at later ages, students may also contribute to the lower propensity to live with a spouse, even though the number of kids is not that different between our sample and the population in general.

These differences do not affect the internal validity of our results, or the mechanism we describe. On the external validity of the magnitudes of some responses, we underestimate the general population's substitution between liquidity and insurance if renters have even stronger consumption commitments or tighter liquidity constraints (if they did not qualify for mortgages).

## 1.4 Results

In this section we present evidence that after 1992 homeowners with much home equity reduced their demand for unemployment insurance compared to owners with large mortgages on their homes, which left them unaffected by the mortgage reform. Because our empirical strategy relies on the common trend assumption across various levels of leverage, we study the correlation between year-on-year changes over time in the proportion of insured and home equity by December 31, 1991, and we present evidence that insurance trends did not differ significantly in the pre-reform period across groups. In subsection 1.4.2, we perform placebo tests on the years before the reform and show that, without the mortgage reform, home equity at time  $t$  has no impact on the demand for unemployment insurance at time  $t + 1$ .

### 1.4.1 Main Results

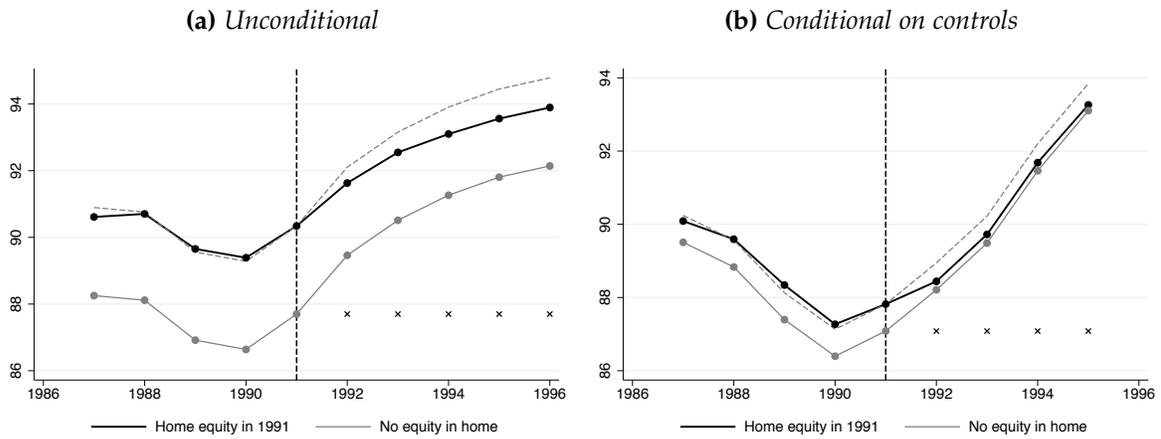
We perform our analysis of the effect of the liquidity shock on unemployment insurance membership rates after 1992 with both a discrete and a continuous formulation for the liquidity treatment. We define the discrete treatment group as those having more than a month of permanent income worth of credit in 1991 home equity. The control group for the discrete formulation are those homeowners whose liquidity shock measure is less than or equal to zero.<sup>18</sup> We compare the trends over time in insurance demand of the two groups in Figure 1.3. The figure plots in both panes the insured share over time in solid black for the treated group and in solid gray for the control group. For easy comparison, the dashed gray line represents the parallel shift of the control group up towards the treated group, such that 1991 percentages coincide for both groups. The figure marks the years in which the difference between the control group average and the projection of the treatment group average is statistically significant at the 1% confidence level according to a proportion test.

Figure 1.3 is divided into two panes. Pane 1.3a plots the unconditional yearly share of

---

<sup>18</sup>To mitigate eventual misclassification bias between the treatment and control groups, we exclude from this analysis individuals with positive home equity worth less than a month of permanent income.

**Figure 1.3: Insurance Up-Take by Treatment Group around the Reform**



NOTE.— The black solid line shows the average insurance up-take over time for those homeowners who experienced a liquidity shock larger than a month’s worth of their permanent income in 1992; the gray solid line indicates the average insurance up-take over time for those homeowners who experienced no liquidity shock, that is they held no equity they could borrow against in December 1991. We define our measure of liquidity shock in equation (1.7). The dashed gray line shifts the average insurance up-take of those who did not experience a liquidity shock such that the average insurance rates of the two groups coincide in 1991. The estimation sample comprises homeowners between ages 29 and 34 in 1991 – for more detailed information about sample selection, see Section 1.3.

insurance over time for the two groups. Pane 1.3b plots the same trends conditioned on a set of controls, which includes marital status, gender, the number of children below the age of 18 in the household, disposable income in the year, unemployment risk, and the 1991 values of liquid assets and debts. Additionally, we control for year, cohort, industry and education fixed effects, and for our measure of permanent income.

Both the unconditional and the conditional analysis show that the trend for the treated group and the projected trend for the control group closely follow each other before the reform, but they significantly diverge from 1992 onwards. That the pre-trends line up closely before 1992 supports the common trends assumption, on which our identification strategy rests. The figure shows that treated individuals reduce their demand for unemployment insurance after the reform relative to the control group. Unconditionally, the effect seems to unfold over time after 1992, trends controlling for unobservables diverge sharply in 1992, and that their difference remains constant afterwards.

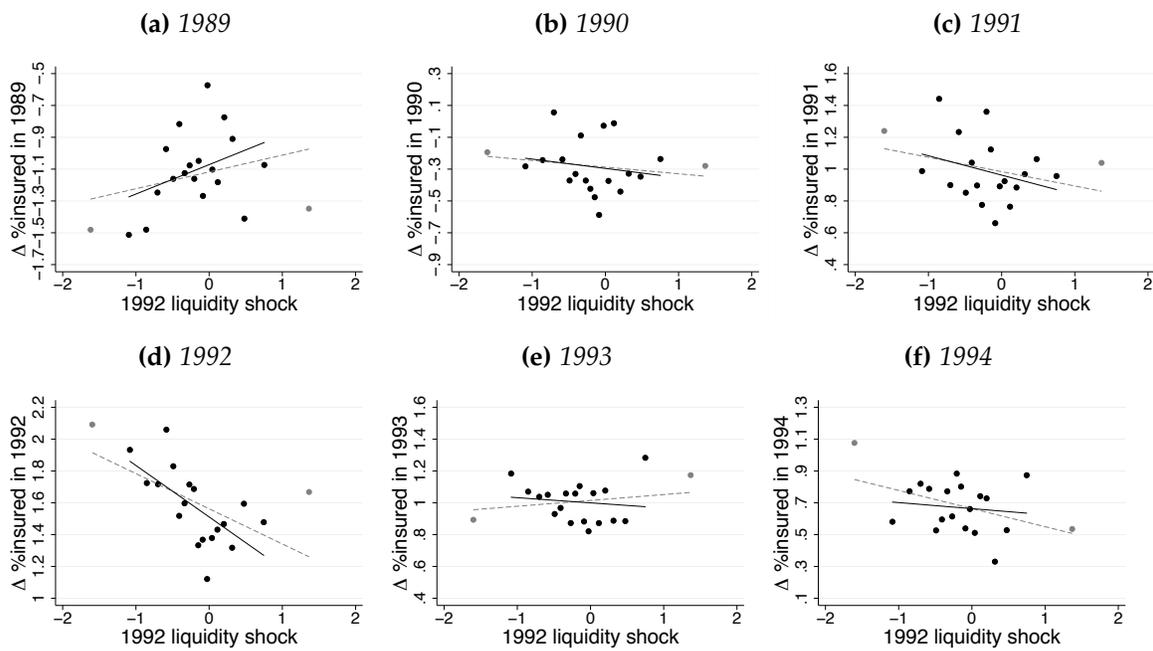
This finding does not depend on the chosen cutoff between the treated and the control groups. Similar results hold across treatment intensities. To study the trends across treatment intensities, we focus on the partial correlation between accessible home equity in 1991 and the yearly (net) changes in insurance purchase, i.e. subscriptions minus unsubscriptions. Figure 1.4 plots yearly conditional changes in insurance purchase, by vigintiles of conditional home equity in 1991.<sup>19</sup> This procedure returns a set of twenty groups with different treatment intensity (or dosage), and a more credible graphical analysis of the partial correlation between treatment and outcome over the years.<sup>20</sup>

---

<sup>19</sup>For each year in our sample, we regress both the first difference in insurance purchase and accessible home equity in 1991 on our set of controls. We divide the residuals from the regression on accessible home equity in 1991 into twenty vigintiles, and for each vigintile we plot in Figure 1.4 the average residuals from the two regressions, adding back the overall mean.

<sup>20</sup>Figure 1.4 shades the top and bottom 5% by treatment intensity. These percentiles include extreme values of the treatment variable, potentially caused by the time difference between property transactions and mortgage contracts being recorded in the registers even after our sample restrictions. As we cannot test this hypothesis, we keep those observations in our analysis for the sake of robustness, but we present two different linear regression lines in Figure 1.4 to show the magnitude of the correlation with and without these vigintiles: The dashed gray lines show regression lines for the full sample; the solid black lines for the sample excluding the top and bottom vigintiles. As expected, errors in the extreme values of our measure of the liquidity shock dilute the effect of the 1992 reform, consistent with the hypothesis that these are for the most part due to measurement error.

**Figure 1.4:** Impact of 1991 Home Equity on Unemployment Insurance Sign-Up (with controls)



NOTE.— This figure plots yearly percentage point changes in insurance up-take, conditional on controls, by vigintiles of home equity in 1991. See footnote 19 for details and Section 1.3 for details of our sample selection. For easy comparison, the scale of the axes is constant throughout the panes. Intercepts differ because of different mean net subscription rates. The dashed gray lines show the regression lines for the full sample; the solid black lines for the sample excluding the top and bottom vigintiles.

Pane 1.4d shows that only in 1992 is the correlation between the changes in the net unemployment insurance membership and the treatment strong and negative: The more home equity individuals hold by the end of 1991, the less insurance they buy in 1992 after the reform. After 1992, the correlation disappears, confirming the post-reform stabilization of the effect shown in Figure 1.3.

Figure 1.4 shows that the negative effect of a shock to liquidity on demand for insurance depends on treatment intensity (or dosage). We report partial treatment effect estimates of the liquidity effect for the continuous treatment definition in Table 1.2, with the same set of controls as in Figures 1.3b and 1.4.

The first column of Table 1.2 collects the estimated coefficients for a linear probability model of insurance purchase, where the first row shows the estimated coefficients for the liquidity shock measure interacted with the post-reform period. These are the partial effects on the probability of buying insurance in percentage points. Standard errors are clustered at the individual level, allowing for arbitrary autocorrelation of errors within individuals across the pre- and post-reform periods (Bertrand *et al.*, 2004).

According to the OLS estimates, potential access to credit equivalent to one year of estimated permanent income decreases the probability of buying unemployment insurance by approximately 0.5 percentage points. The estimate is highly significant, and robust to individual fixed effects, as reported in the second column of the table. Because unemployment insurance in Denmark is subsidized and convenient for all those who face non-trivial unemployment risk, and because we do not observe the other adjustment channels that increased access to liquidity crowds out (e.g. durable consumption as in Browning and Crossley, 2009), we interpret this estimate as a lower bound on the effect such a reform would have in an environment with fewer rigidities, more salience of the insurance decision, and fewer people facing better than actuarially fair prices. This finding suggests that liquidity affects demand for insurance significantly, inducing people on the margin of insurance choice to change their behavior.

As insurance choice is affected by many unobserved idiosyncrasies (e.g. risks, cir-

**Table 1.2:** *Impact of 1992 Mortgage Reform on Unemployment Insurance Participation (treatment with dosage)*

	OLS	FE	LDV	FE Logit	LDV Logit
1991 home equity, after 1991	-0.508** (0.0835)	-0.460** (0.0826)	-0.281** (0.0410)	-0.775 (0.401)	-0.130** (0.0327)
1991 home equity	0.825** (0.114)		0.239** (0.0331)		0.119** (0.0249)
1991 liquid assets	0.459* (0.226)		0.148** (0.0426)		0.120** (0.0301)
Permanent income	-0.615** (0.0299)		-0.104** (0.00769)		-0.0424** (0.00415)
1991 debt	-2.100** (0.186)		-0.424** (0.0367)		-0.274** (0.0336)
1991 housing wealth	-1.109** (0.0920)		-0.179** (0.0206)		-0.165** (0.0188)
Disposable income	-10.30** (0.986)	0.867 (0.636)	-1.768** (0.267)	5.058* (2.216)	-0.125 (0.162)
Unemployment risk (pp.)	0.817** (0.0251)	0.319** (0.0187)	0.0725** (0.00648)	1.386** (0.0930)	0.0510** (0.00555)
Experience (year)	-1.802** (0.0811)	-1.129** (0.0656)	-0.229** (0.0242)	-9.541** (0.504)	-0.281** (0.0285)
Number of kids	-0.377** (0.0746)	0.117 (0.0611)	-0.141** (0.0176)	1.745** (0.332)	-0.129** (0.0166)
Female	2.949** (0.200)		0.614** (0.0427)		0.451** (0.0354)
Married or cohabiting	1.124** (0.133)	0.638** (0.106)	0.247** (0.0335)	5.299** (0.587)	0.269** (0.0312)
Lagged insurance			84.44** (0.125)		10.84** (0.0399)
Year	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	No	Yes	No	Yes
Industry	Yes	Yes	Yes	Yes	Yes
Education	Yes	No	Yes	No	Yes
Municipality	Yes	Yes	Yes	Yes	Yes
Income vigintiles	Yes	Yes	Yes	Yes	Yes
Observations	1,020,096	1,020,096	906,752	130,986	906,752
Individuals	113,344	113,344	113,344	14,554	113,344

NOTE.—Standard errors clustered by individual in parentheses; \*  $p < 0.05$ , \*\*  $p < 0.01$ . The table collects Average Partial Effects (APE) on insurance up-take computed for the post-1991 subsample, in percentage points, from models of 0-1 unemployment insurance status estimated on the entire 1987-1995 panel. These models include the continuous measure of home equity (relative to permanent income) in 1991 as the dosage of the extra liquidity treatment afterwards. The first column shows the estimates from an OLS model. The second column includes individual fixed effects. The model in the third column adds the lagged insurance status to control for inertia. The fourth column shows the APE from a fixed effect logit model, estimated on only those who ever change their insurance status. Finally, the model in the fifth column is a discrete choice model with fixed costs of switching, which translates into a lagged dependent variable logit. Financial variables are scaled by annual permanent income. The estimation sample is the fully balanced 1987-1995 panel of homeowners between ages 29 and 34 in 1991—for more detailed information about sample selection, see Section 1.3.

cumstances, preferences), it is natural to extend the model with individual fixed effects.<sup>21</sup> Column 2 shows that while the explanatory power of the model rises considerably across specifications, the coefficient of interest changes slightly. This result allows us to apply the bounding exercise introduced by Oster (2014) to calculate what the true treatment effect could be in the linear model under an assumption of proportional selection on unobservables (relative to observables, including the fixed effects). We compare the specification including individual fixed effects with an unconditional difference-in-differences model that does not control for observables, and we assume that the confounding effects of observable variables are proportional to those of unobservable confounders. The identification of the treatment effect in the unconditional difference-in-differences design requires the inclusion of only the group identifier (or in this continuous specification, end-of-91 home equity) and year dummies as controls. This (unreported) model has a point estimate of -0.544 for the treatment effect, with an  $R^2$  of only 0.0046. In the conditional model that includes fixed effects (column 2), the point estimate is -0.460, while the  $R^2$  of the model is 0.7416. This finding suggests that the true treatment effect is larger in magnitude than  $-0.460 - [-0.544 - (-0.460)] \frac{1-0.7416}{0.7416-0.0046} = -0.43$  percentage points of insurance purchase.<sup>22</sup>

The subsidies to the insurance system in Denmark and the social norms associated with unemployment fund membership (see section 1.2) imply that the majority of the population always insures, and even others rarely change membership status. As described in Section 1.2, we model inertia in two alternative random utility models. In the fourth column, we show the estimates from a fixed-effect logit model, which estimates the parameters of the model on the 13% subsample that changes insurance membership status during the period of interest. While we estimate a negative coefficient associated with the liquidity shock

---

<sup>21</sup>Fixed effects cannot confound our estimates identified from differential changes after 1992, yet they improve the precision of the treatment effect estimates and controls.

<sup>22</sup>This calculation assumes that the unobservable confounders are not only proportional, but also equally powerful explanatory factors, i.e.  $\delta = 1$  in the notation of Oster (2014). Specification errors or invalid identification clearly remain a threat to the estimation of this effect. The nonlinear specifications and placebo tests that follow are thus complementary to this argument.

provided by the reform in this model, this estimate is just below a 5% significance threshold.

In column 5, we show the predicted partial effect of liquidity (more home equity after the reform) from the random utility model with fixed switching costs for changing membership status described in equation (1.6). This model explicitly incorporates inertia in the estimation. As a comparison, we show the coefficient associated with its linear probability model counterpart in column 3. The results from the model in column 5 show a significant negative effect associated with the liquidity shock after 1991, and are thus in line with those from the linear probability models. However, this model also highlights the importance of inertia and state dependence in our setup. The sum of the costs for changing unemployment fund membership accounts for over six times as much of the variation in insurance decisions as the unobservable variation in the latent model.

Across the columns of Table 1.2, we can compare how the predicted partial effect of liquidity (more home equity after the reform) changes across the specifications.<sup>23</sup> The linear model predicts that 1 in 200 Danes choose not to buy formal unemployment insurance because of the extra liquidity. This estimate is robust to linear fixed effects but, as many of the observations have high baseline probability to insure, is smaller when computed according to the nonlinear models. The effects seem half as strong in the linear model controlling for previous insurance status, and a fourth as large in a logit model with the same control for persistence. The few observations that allow a fixed effects logit estimation have a larger point estimate, but 5-10 times the standard errors. Covariates have effects with the expected signs, e.g. the risk of unemployment significantly raises the chance of buying unemployment insurance (roughly 1-to-1 in percentage points). The strong persistence of the insurance decision is also evident, indicating large inertia in insurance choices.

We repeat our baseline estimation and robustness checks for the discrete treatment definition in Table 1.3. While a discrete treatment allows for a more straightforward

---

<sup>23</sup>This prediction calculates the marginal effects at the observed levels of all covariates in the post-1991 period. This measure is closer to the ATET (average treatment effect on the treated) than that obtained using the full sample. Because overall insurance up-take increased after 1992, we expect smaller marginal effects using the post-1991 period than using the full sample.

implementation of the difference-in-differences estimator, we lose information relative to the continuous treatment definition. We still find a highly significant effect with the OLS and fixed effects estimators.<sup>24</sup> Homeowners who gained access to extra liquidity in 1992 decreased their likelihood of purchasing unemployment insurance by 0.7 percentage points compared to homeowners who already mortgaged to the limit. Our results using the random utility models are similar: Partial effects in the linear specification suggest a roughly half a percentage point drop in insurance up-take because of the liquidity buffer after the reform. However, the predicted drop in insurance probability is half as much once we control for the persistence of the insurance decision, and only -0.2 if we do so in a logit model. This finding suggests that the insurance choice is closely related to the amount of accessible credit, rather than access to credit itself.

#### **1.4.2 Placebo Tests and Robustness Checks**

Our results indicate that those who gained access to home equity due to the 1992 reform decreased their demand for unemployment insurance compared to those who had no access. In this section, we address two potential mechanisms that might confound our results. First, our treatment selects people with much home equity by December 31, 1991, and we cannot a priori distinguish the effect of the reform from the effect of having much home equity in one given year.

Second, home purchases and mortgaging decisions are strongly dependent on household formation choices, and household composition itself affects the attractiveness of the unemployment insurance scheme. In particular, the alternative of supplemental security income changes with marriage, as the means testing for supplemental security income is more severe for couples. We tackle these two concerns separately.

To identify the specific effect of the 1992 reform, and rule out that the effect in Table 1.2 is in fact caused by mechanical correlates of treatment, we repeat our analysis for a series of

---

<sup>24</sup>As individual fixed effects again raise the  $R^2$  to 0.742, the upper bound on the treatment effect is -0.48 percentage points under the necessary assumptions for applying Oster (2014), which we state earlier in the text. We conclude that omitted variable bias does not seem to threaten our finding of a small but significant effect.

**Table 1.3:** *Impact of 1992 Mortgage Reform on Unemployment Insurance Participation (discrete treatment)*

	OLS	FE	LDV	FE Logit	LDV Logit
Treatment	-0.613** (0.124)	-0.523** (0.123)	-0.364** (0.0605)	-0.306 (0.682)	-0.196** (0.0533)
Treated group	0.615** (0.169)		0.261** (0.0491)		0.122** (0.0405)
1991 liquid assets	0.477* (0.227)		0.156** (0.0427)		0.124** (0.0305)
Permanent income	-0.612** (0.0309)		-0.103** (0.00784)		-0.0415** (0.00423)
1991 debt	-2.034** (0.187)		-0.413** (0.0369)		-0.272** (0.0339)
1991 housing wealth	-1.091** (0.0949)		-0.172** (0.0212)		-0.160** (0.0195)
Disposable income	-10.29** (1.016)	0.685 (0.646)	-1.794** (0.270)	4.374 (2.310)	-0.128 (0.164)
Unemployment risk (pp.)	0.831** (0.0260)	0.319** (0.0192)	0.0734** (0.00669)	1.383** (0.0954)	0.0522** (0.00573)
Experience (year)	-1.780** (0.0840)	-1.144** (0.0680)	-0.229** (0.0250)	-9.661** (0.488)	-0.283** (0.0295)
Number of kids	-0.351** (0.0771)	0.131* (0.0633)	-0.138** (0.0182)	1.868** (0.344)	-0.127** (0.0172)
Female	2.923** (0.207)		0.616** (0.0441)		0.454** (0.0366)
Married or cohabiting	1.078** (0.137)	0.641** (0.109)	0.234** (0.0346)	5.341** (0.604)	0.254** (0.0323)
Lagged insurance			84.47** (0.128)		10.91** (0.0412)
Year	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	No	Yes	No	Yes
Industry	Yes	Yes	Yes	Yes	Yes
Education	Yes	No	Yes	No	Yes
Municipality	Yes	Yes	Yes	Yes	Yes
Income vigintiles	Yes	Yes	Yes	Yes	Yes
Observations	1,020,096	1,020,096	906,752	130,986	906,752
Individuals	113,344	113,344	113,344	14,554	113,344

NOTE.—Standard errors clustered by individual in parentheses; \*  $p < 0.05$ , \*\*  $p < 0.01$ . The table collects Average Partial Effects (APE) on insurance up-take computed for the post-1991 subsample, in percentage points, from models of 0-1 insurance status estimated on the entire 1987-1995 panel. These models include a discrete measure of any home equity vs none in 1991 as a binary treatment of any extra liquidity afterwards. Any home equity here is more than a month's income below the LTV limit, while none corresponds to being over the LTV limit. The first column shows the estimates from an OLS model. The second column includes individual fixed effects. The model in the third column adds the lagged insurance status to control for inertia. The fourth column shows the APE from a fixed effect logit model, estimated on only those who ever change their insurance status. Finally, the model in the fifth column is a discrete choice model with fixed costs of switching, which translates into a lagged dependent variable logit. Financial variables are scaled by annual permanent income. The estimation sample is the fully balanced 1987-1995 panel of homeowners between ages 29 and 34 in 1991 — for more detailed information about sample selection, see Section 1.3.

placebo reforms, taking place in all years in our sample before 1992. Figure 1.5 shows the equivalent of Figure 1.4 for placebo reforms from 1988 to 1991. We plot partial correlations between the net insurance sign-up in a given year and the amount of home equity in the year before, using the same regressors and specifications used for the analysis of the 1992 reform.<sup>25</sup> That is, in the first panel we plot the change in the percentage of insured between 1987 and 1988 on the vigintiles of unexplained 1987 home equity. The sample selection for the placebo analyses carries over from the 1992 reform, except that for each placebo year, we only keep homeowners in the year preceding the placebo reform. Therefore, the number of observations changes year by year.

All placebo tests exhibit unemployment insurance–home equity correlations that scatter around zero, with no systematic pattern that Figure 1.4d would fit into. This finding not only supports the validity of our controls, but it also rules out that home equity has a mechanical effect on demand for insurance, independently of liquidity, and therefore supports the causal interpretation of our estimates.

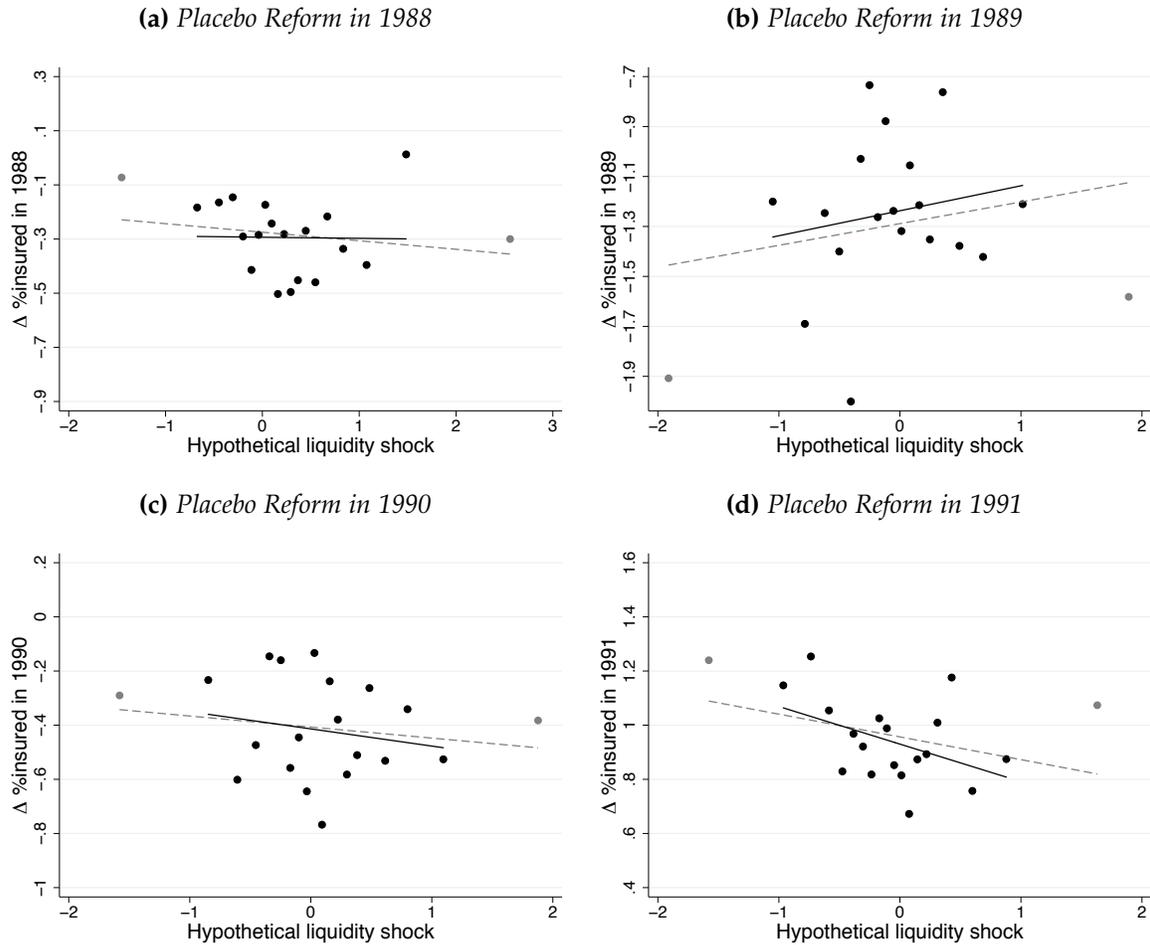
As a second earner can also cushion shocks and even affects whether one qualifies for the fallback benefits, our results can be confounded if individuals with more home equity also form more households. To rule out this confounding channel, we estimate our models using a specific subsample of stable households. Instead of controlling for household size (single or couple), as we did in our baseline specification, we estimate the models in Tables 1.2 and 1.3 after excluding observations for which marital status is different from that of 1991. That is, if we observe an individual getting married in 1989, we keep only observations from 1989 onwards. This way we obtain a subsample of observations that, though unbalanced, contains only households with stable marital status throughout the estimation period.

Table 1.4 shows that results are robust to restricting the sample to stable households. All controls and model specifications are the same as those in Tables 1.2 and 1.3. Compared to our baseline estimates, these results are similar, if not stronger. We therefore argue that the

---

<sup>25</sup>We hold wealth controls constant at their pre-reform levels. That is, for each of the (placebo) reform years, we control for wealth held by the end of the previous year.

**Figure 1.5:** *Placebo Tests: Impact of Pre-1991 Home Equity on Unemployment Insurance Sign-Up (with controls)*



NOTE.—The figure shows the analogue of Figure 1.4 Panel (d) for placebo reforms from 1988 to 1991. Thus it plots yearly percentage point changes in insurance purchase, conditional on controls, by vigintiles of home equity in the year of the placebo shock. See footnote 19 for details and Section 1.3 for details of our sample selection.

**Table 1.4:** *Impact of 1992 Mortgage Reform on Unemployment Insurance Participation among Households Stable around 1991*

	OLS	FE	LDV	FE Logit	LDV Logit
Continuous	-0.611** (0.103)	-0.516** (0.101)	-0.315** (0.0504)	-1.045** (0.405)	-0.170** (0.0454)
Observations (C)	622,521	622,521	553,352	77,598	552,915
Individuals (C)	69,169	69,169	69,169	8,622	69,114
Discrete	-0.804** (0.158)	-0.628** (0.156)	-0.373** (0.0771)	-1.178 (0.846)	-0.215** (0.0773)
Observations (D)	587,970	587,970	522,640	73,818	522,234
Individuals (D)	65,330	65,330	65,330	8,202	65,279

NOTE.—Standard errors clustered by individual in parentheses; \*  $p < 0.05$ , \*\*  $p < 0.01$ . The table collects Average Partial Effects (APE) on insurance up-take computed for the post-1991 subsample, in percentage points, from models of 0-1 unemployment insurance status estimated on the entire 1987-1995 panel. The continuous and discrete specifications correspond to Tables 1.2 and 1.3, respectively, restricted to those whose marital status does not change over the period. See Section 1.3 for details of our sample selection otherwise. The first column shows the estimates from an OLS model. The second column includes individual fixed effects. The model in the third column adds the lagged insurance status to control for inertia. The fourth column shows the coefficient in a fixed effect logit model, estimated on only those who ever change their insurance status. Finally, the model in the fifth column is a discrete choice model with fixed costs of switching, which translates into a lagged dependent variable logit. The estimation sample is the 1987-1995 panel of homeowners between ages 29 and 34 in 1991, when they live in the same household as in 1991 — for more detailed information about sample selection, see Section 1.3. All models include the same controls used for the estimates shown in Tables 1.2 and 1.3.

results in Table 1.2 and 1.3 are not driven by differential patterns in household formation across levels of home equity.

### 1.4.3 Heterogenous Effects

Figure 1.1 showed that we predict insurance to have positive expected value for a large fraction of the population in 1987, and this is true for all years. If these people are making a fully-informed rational decision about insurance, liquidity should be irrelevant for them; they should buy insurance regardless. The average treatment effect in the population is supposed to come from the left tail of the risk distribution. In Table 1.5, we show our main specification over five equal-sized cuts of the 1992 risk distribution, and indeed, for the lowest quintile with average risk of 1.85% in 1992, the estimated effect is double the population average from the corresponding column 1 of Table 1.2. Relaxing liquidity constraints by one year's worth of permanent income decreases insurance purchase by 0.94 percentage points. To put this magnitude into context, this amounts to more than 5% of this low-risk subpopulation who was without insurance in 1991, though some of the effect also comes from insured Danes not renewing their membership. For higher risk quintiles, where insurance is a bargain, the estimated effect of more liquidity is not only lower, but not significantly different from zero.

We also argued that the insurance system faces some rigidities, and it might be puzzling why not all high-risk Danes join a fund, or why low-risk occupations are ready to cross-subsidize others. Yet for those making an active choice about insurance, we can compare two competing incentives: How does the effect of liquidity compare to that of a 1 percentage point reduction in unemployment risk? From the point estimates for the low-risk quintile in Table 1.5, we can conclude that one year's income in liquidity has similar effects as a 0.3 percentage point drop in unemployment risk, or 15% of their baseline risk in 1992, among those for whom insurance is priced most unfairly.<sup>26</sup>

---

<sup>26</sup>For the correct interpretation of this rescaling, we do not claim to have identified the causal impact of unemployment risk, nor that our predicted risk measure is an unbiased and properly scaled estimate of subjective risk perceptions each year.

**Table 1.5:** *Impact of 1992 Mortgage Reform on Unemployment Insurance Participation by Unemployment Risk Quintiles*

	Risk Q1	Risk Q2	Risk Q3	Risk Q4	Risk Q5
1991 home equity, after 1991	-0.944** (0.245)	-0.440* (0.184)	-0.128 (0.200)	-0.0542 (0.136)	-0.275 (0.151)
1991 home equity	1.146** (0.300)	0.508 (0.272)	2.059** (0.309)	0.200 (0.186)	0.112 (0.218)
1991 liquid assets	0.974 (0.588)	0.517 (0.572)	0.307 (0.494)	0.212 (0.392)	-0.620 (0.515)
Permanent income	-0.864** (0.0580)	-0.627** (0.0591)	-0.751** (0.0725)	-0.363** (0.0613)	-0.360** (0.0738)
1991 debt	-3.424** (0.517)	-2.583** (0.441)	-1.857** (0.403)	-1.288** (0.354)	-1.120** (0.341)
1991 housing wealth	-1.398** (0.244)	-1.081** (0.221)	-1.808** (0.271)	-0.670** (0.150)	-0.830** (0.163)
Disposable income	-14.37** (1.988)	-8.406** (2.043)	-5.854* (2.471)	-4.405* (1.846)	-3.822 (2.018)
Unemployment risk (pp.)	3.298** (0.171)	0.193 (0.121)	0.0539 (0.114)	-0.287** (0.0873)	-0.0861 (0.0471)
Experience	-2.526** (0.189)	-0.807** (0.243)	-3.663** (0.234)	-0.290 (0.158)	-0.968** (0.141)
Number of kids	-0.328 (0.189)	-0.182 (0.173)	-0.626** (0.215)	-0.200 (0.120)	-0.268* (0.126)
Female	5.161** (0.499)	0.122 (0.475)	3.052** (0.472)	1.155** (0.333)	-0.248 (0.396)
Married or cohabiting	1.374** (0.326)	1.690** (0.315)	0.126 (0.377)	1.066** (0.212)	1.075** (0.227)
Year	Yes	Yes	Yes	Yes	Yes
Cohort	Yes	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes	Yes
Education	Yes	Yes	Yes	Yes	Yes
Municipality	Yes	Yes	Yes	Yes	Yes
Income vigintiles	Yes	Yes	Yes	Yes	Yes
Observations	219,319	188,580	189,880	20,1792	196,009
Individuals	24,401	20,967	21,131	22,462	21,822

NOTE.—Standard errors clustered by individual in parentheses; \*  $p < 0.05$ , \*\*  $p < 0.01$ . The table reports the estimated coefficient associated with our treatment variable in the continuous specification, by five quintiles of 1992 unemployment risk. These estimates correspond to the OLS estimates in column 1 in Table 1.2. The table collects coefficients from OLS models of 0-1 insurance status. Financial variables are scaled by annual permanent income. The estimation sample comprises homeowners between 29 and 34 in 1991 — for more detailed information about sample selection, see Section 1.3. Unemployment risk here is the average FTFY equivalent time spent on benefits in 1993 for others in the estimation sample who are full-time insured in the same industry and broad education category in November 1992. This leave-out mean unemployment risk predicts realized unemployment with an  $R^2$  of 0.59 over the 1987-1995 period.

## 1.5 Conclusion

If liquidity is a pressing concern during unemployment, people will be partially protected by a buffer stock of savings. This paper documents how increased access to liquidity through the exogenous introduction of home equity loans lowered the demand for unemployment insurance in Denmark, implying that private self-insurance substitutes for formal public insurance. The demand for unemployment insurance increased in Denmark throughout our period of interest. However, after the 1992 reform, demand increased relatively less for those who held equity in their homes compared to those who did not. By exploiting the unique policy-induced variation provided by a mortgage reform, we show that access to liquidity affects insurance choices on the margin, even when wealth does not change with it. Simply relaxing liquidity constraints shields people from misfortune to such an extent that some prefer to avoid paying an unemployment insurance premium.

We show that an additional increase in accessible liquidity worth one year of income caused only about 0.5 percentage of Danes to forgo public unemployment insurance. Among individuals for whom insurance is much more expensive than actuarially fair, a year's income's worth of extra liquidity reduces insurance up-take by 0.94 percentage points. This effect is equivalent to that of a 0.3 percentage point, or 15%, decrease in the risk of unemployment, while higher-risk groups show no effect.

Our findings relate to the discussion about the scope of social insurance programs and whether unemployment insurance should be mandatory: The mere option to use one's own resources more flexibly alleviates the welfare costs from job loss. Some workers in our sample were able to perceive this opportunity in a forward-looking manner and to make a conscious insurance choice accordingly. While the modest crowd-out reminds us of other important drivers of insurance up-take, this finding suggests that increased access to liquidity for the general population substitutes partially for a publicly funded unemployment insurance scheme.

## Chapter 2

# What Policies Increase Prosocial Behavior? An Experiment with Referees at the *Journal of Public Economics*<sup>1</sup>

### 2.1 Introduction

The peer review process familiar to all academic researchers offers a classic example of the positive externalities from prosocial behavior: the reviewer bears the costs from submitting a high-quality referee report quickly, while the gains to the authors of the paper and to society from the knowledge produced are potentially large. We evaluate the impacts of economic and social incentives on peer review using an experiment with 1,500 referees at the *Journal of Public Economics*. The specific aim of the experiment is to understand how to improve the speed and quality of peer review, an issue of particular importance to the economics profession given the slowdown of the publishing process (Ellison, 2002). Our

---

<sup>1</sup>Co-authored with Raj Chetty and Emmanuel Saez

broader objective is to evaluate commonly used methods of increasing prosocial behavior and to test the predictions of competing theories.

In our experiment, we randomly assign referees to four groups: a control group with a six-week (45 day) deadline to submit a referee report, a group with a four week (28 day) deadline, a cash incentive group rewarded with \$100 for meeting a four week deadline, and a social incentive group in which referees were told that their turnaround times would be publicly posted. The experiment yields four sets of results.

First, shortening the deadline from 6 weeks to 4 weeks reduces median review times from 48 days to 36 days. Because missing the deadline has no direct consequence, we believe the shorter deadline acts primarily as a “nudge” (Thaler and Sunstein, 2008) that changes the default date at which referees submit reports. Second, providing a \$100 cash incentive for submitting a report within four weeks reduces median review times by an additional eight days. Third, the social incentive treatment reduces median review times by approximately 2.5 days—which is intriguing given that the degree of social pressure applied here is relatively light. We also find that that social incentives have much larger effects on tenured professors, but in contrast, tenured professors are less sensitive to deadlines and cash incentives than untenured referees.

Finally, we evaluate whether the treatments have an impact on other outcomes besides review time.<sup>2</sup> Economic models of multi-tasking (e.g., Holmstrom and Milgrom, 1991) predict that referees will prioritize the incentivized task (i.e., submitting a report quickly) at the expense of other aspects of performance (e.g., the quality of reviews). We find that the shorter deadline has no effect on the quality of the reports that referees submit, as measured by whether the editor follows their recommendation or the length of referee reports. The cash and social incentives induce referees to write slightly shorter referee reports, but do not affect the probability that the editor follows the referee’s advice. We also find little evidence

---

<sup>2</sup>The cash incentive increases the fraction of referees who agree to review a manuscript. The social incentive reduces agreement rates, while the shorter deadline has no impact. We show that the selection effects induced by these changes in agreement rates are modest and are unlikely to explain the observed changes in review times.

of negative spillovers across journals: the treatments have no detectable effects on referees' willingness to review manuscripts and review times at other Elsevier journals.

We conclude that small changes in journals' policies could substantially improve the peer review process at little cost. Shorter deadlines appear to be an essentially costless means of expediting reviews. Cash and social incentives are also effective, but have monetary and psychic costs that must be weighed against their benefits.

A large body of evidence from the lab has considered the determinants of prosocial behavior and altruism (for example, Ledyard, 1995; Fehr and Fischbacher, 2003; Vesterlund, 2012). Our study provides evidence from the field, which has been considerably more limited. Prior work concerning prosocial behavior has often debated whether extrinsic incentives such as cash payments are effective in increasing prosocial behavior because they may crowd out intrinsic motivation (Titmuss, 1971; Bénabou and Tirole, 2006). In our application, if referees submit reviews to be recognized for their service to the profession by editors, the provision of monetary incentives could potentially erode this signal and have a negative impact on review times. However, our analysis shows that, at least in this context, price incentives, nudges, and social pressure are all effective and complementary methods of increasing prosocial behavior.

## **2.2 Experimental Design**

We conducted the experiment over a 20 month period, from February 15, 2010 to October 26, 2011. All referees for the *Journal of Public Economics* during this period were randomly assigned to one of four groups. For simplicity, only referee requests for new submissions were included in the experiment. These assignments were permanent for the duration of the experiment: referees never switched groups. The co-editors in charge of handling each new submission chose referees to review the paper without seeing the group to which the referee was assigned.

**Table 2.1:** *Description of Treatment Groups*

Group:	6 Week	Social	4 Week	Cash
Deadline	6 weeks (45 days)	6 weeks (45 days)	4 weeks (28 days)	4 weeks (28 days)
Incentives	None	Review time posted online at end of year	None	\$100 Amazon gift card if deadline met
Duration of Intervention	Feb 15, 2010 to Oct 26, 2011	Feb 15, 2010 to Oct 26, 2011	Feb 15, 2010 to Oct 26, 2011	Feb 15, 2010 to May 9, 2011

NOTE.—This table describes the four treatment groups to which referees were randomly assigned. Every referee was assigned permanently to one group; referees never changed groups. Referees were notified about the conditions of the review request upon invitation and were sent a reminder 1 week before the deadline. Examples of these invitation and reminder emails are shown in Appendices A and B. Cash incentives were stopped for invitations after May 9, 2011; after that point, referees assigned to the cash incentive group simply faced a 4 week deadline, with no incentives. The other treatments were implemented without any changes for the full duration of the experiment, from February 15, 2010 to October 26, 2011.

Some key features of the four groups are shown in Table 2.1.<sup>3</sup> All deadlines for the differing groups were defined relative to the date at which the invitation was sent—not the date at which the referee accepted the invitation—to eliminate incentives to delay agreement.

The *control* or what we will refer to as the *six-week* group actually faced a 45 day deadline for submitting a referee report, the deadline that was in place at the journal before the experiment began. The deadline was described using the following language in the invitation letter: “If you accept this invitation, I would be very grateful if you would return your review on or before **July 21, 2010** (6 weeks from now).”

The *four-week* group faced a 28 day deadline for submitting a report. The email they received was identical to that sent to referees in the control group, except for the due date.

The *cash incentive* group faced a 28 day deadline and received a \$100 Amazon gift card

---

<sup>3</sup>An appendix includes the details of the experiment. Figure B.1 presents a flow chart for the entire experiment. Appendix B.7 shows our invitation emails. Appendix B.8 shows our reminder and thank-you emails. Appendix B.3 includes more detail on data sources and variable definitions. Appendix Table B.1 presents summary statistics for the primary experimental period (referee invitations between February 15, 2010 and May 9, 2011). Appendix B.4 describes the reweighting methodology behind Figure 2.2b. Appendix Table B.4 presents the hazard model estimates of treatment effects on review times. Appendix B.6 provides a list of other journals used to assess spillover effects. Appendix B.9 presents a summary of all the appendix tables and figures. A de-identified version of the 3,397 observation dataset is available at [http://obs.rc.fas.harvard.edu/chetty/jpube\\_experiment.zip](http://obs.rc.fas.harvard.edu/chetty/jpube_experiment.zip).

for submitting a report before the deadline. In addition to the standard text describing the deadline, the invitation letters in the cash incentive group included the following text: “As a token of appreciation for timely reviews, you will receive a \$100 Amazon.com® Gift Card if you submit your report on or before the due date. The Journal of Public Economics will automatically email you a gift card code within a day after we get your report (no paperwork required).”

Finally, the *social incentive* group faced a six-week (45 day) deadline and was told that referee times would be publicly posted by name at the end of the calendar year. In addition to the standard text describing the deadline, the invitation letters in the social incentive group included the following text: “In the interest of improving transparency and efficiency in the review process, Elsevier will **publish referee times by referee name**, as currently done by the Journal of Financial Economics at this website: [jfe.rochester.edu/colab.htm](http://jfe.rochester.edu/colab.htm). The referee times for reports received in 2010 will be posted on the *Journal of Public Economics* website in January 2011. Note that referee anonymity will be preserved as authors only know the total time from submission to decision (and not individual referee’s times).”

One week prior to their deadlines, referees who had not yet submitted reports received emails reminding them that their reports were due in a week. For the social and cash incentive groups, these emails included language reminding referees of the treatments they faced. We also sent overdue reminders 5 days, 19 days, and 33 days after the due date. Referees in the cash, four-week, and six-week groups were simply informed their reports were past due. Referees in the social incentive group were again reminded that their referee times would be publicly posted. After the referees submitted reports, they received a thank you email. Referees in the cash incentive group received an Amazon gift card code in this thank you email if they submitted before the 28 day deadline. Those in the social incentive group received information on the number of days it took for them to submit the report.

To study the impact of monetary payments on intrinsic motivation after cash incentives are withdrawn, we stopped cash payments on May 9, 2011, roughly six months before we ended the other treatments. Referees in the cash incentive group continued to face a

four-week deadline after this point, and received the same invitation and reminder emails as those in the four-week group. All other treatments continued until the end of the experiment on October 26, 2011, at which point all referees were reverted back to the six-week (45 day) deadline.

We analyze the effects of the experiment using information from two sources. We obtain information on referee assignments, review times, and other related outcomes at the *Journal of Public Economics*, as well as other Elsevier journals from Elsevier's editorial database. We obtain information on referee characteristics—an indicator for holding an academic position, tenure status, gender, and an indicator for working in the United States—from curricula vitae posted online.

Each observation in our analysis dataset corresponds to a single referee invitation sent between February 15, 2010 and October 26, 2011. During this period, 3,397 invitations were sent out to 2,061 distinct referees. We include all observations in the referee report level dataset in our analysis, so that referees who are invited multiple times contribute multiple observations.

In our baseline analysis, we restrict attention to referee invitations sent between February 15, 2010 and May 9, 2011, the period when the cash reward was offered. We term this period the primary experimental period. During this period we sent 2,423 invitations, of which 66.2 percent were accepted. Among these referees, 93.7 percent submitted a report before the editor made a decision. The median turnaround time for those who submitted reports was 41.0 days. Among the 1,157 referees who agreed to review a manuscript during the primary experimental period, 74.9 percent of referees agreed to review one manuscript during the experiment, 16.4 percent agreed to review two manuscripts, and the rest agreed to review three or more manuscripts.

To verify the validity of our experimental design, we calculated these summary statistics by treatment group for referee assignments from November 1, 2005 to February 15, 2010, *before* the experiment began. As expected, given randomization, we find no statistically significant differences across the control group or the three treatment groups in these

pre-determined characteristics (details in Appendix Table B.2). Hence, differences in performance across the four groups during the experimental period can be interpreted as causal effects of the treatments.

## 2.3 Four Sets of Outcomes

We analyze four sets of outcomes: 1) agreement to submit a review, 2) time taken to submit the review, 3) report quality, and 4) performance at other journals.

### 2.3.1 Outcome 1: Acceptance of Referee Invitation

Table 2.2 shows the percentage of referee invitations accepted by treatment group. We structure this and all subsequent tables as follows. The four columns correspond to the four experimental groups: six-week, social, four-week, and cash. For each group, we report the point estimate and associated standard error in parentheses. We cluster standard errors by referee to account for the fact that some referees review multiple papers. We also report p-values for the null hypothesis that agreement rates are the same in each treatment group and its corresponding control group. For the social incentive and four-week deadline groups, the control group is defined as the six-week deadline group. For the cash incentive group, the control group is defined as the four-week deadline group, which is the relevant comparison because the cash incentive group also faced a four-week deadline.

Table 2.2 shows that 67.6 percent of the referee invitations are accepted in the six-week group. The acceptance rate is slightly lower at 61.1 percent in the social incentive group, a difference that is marginally statistically significant ( $p = 0.045$ ). The acceptance rate in the four-week deadline group is 64.1 percent, not significantly different from the acceptance rate in the six-week group. Lastly, the acceptance rate in the cash incentive group is 72.0 percent, which is significantly higher than the acceptance rate in the four-week deadline group ( $p = 0.010$ ).

Consistent with this statistical evidence, the journal received a few emails showing that the treatments influenced the decisions by some referees to review papers. For example,

**Table 2.2:** *Fraction of Referees Who Accept Review Invitation by Treatment Group*

Group:	6 Week	Social	4 Week	Cash
Percent who accept invitation	67.6%	61.1%	64.1%	72.0%
	(2.14)	(2.43)	(2.23)	(2.17)
p-value for equality with control		0.045	0.252	0.010
Observations	639	568	626	590

NOTE.—This table shows the percentage of referees who accept invitations to review in each treatment group. We restrict the sample to invitations sent between February 15, 2010 and May 9, 2011, the time period when the cash reward was offered. Standard errors, clustered at the referee level, are reported in parentheses. We also report p-values for the null hypothesis that agreement rates are the same in each treatment group and its corresponding control group. For the social and 4 week groups, the control group is defined as the 6 week deadline group. For the cash incentive group, the control group is defined as the 4 week deadline group, which is the relevant comparison because the cash incentive group also faced a 4 week deadline. The number of observations (referee report invitations) is reported in the last row.

a referee assigned to the social incentive group wrote, “I was surprised to receive an email stating the journal is posting referee times by names. . . I would like to withdraw my agreement to referee this paper. Sorry about that. I would have been happy to send in a report on time under a different policy.” Other referees’ emails explain why cash incentives increase acceptance rates. For instance, a referee in the control group wrote, “I am sorry to have to decline this “invitation” to work for free. . . Can’t Elsevier offer a better reward for the time they ask to devote to this screening?” Overall, these results allay the concern that pushing referees to submit reviews quickly will make it difficult to find referees who are willing to submit reviews.

## 2.4 Outcome 2: Review Time

We now turn to the central outcome our treatments were designed to change: the time that referees take to submit their reviews. Naturally, we can only observe review times for referees who agree to submit reviews. Because the referees who accept invitations may differ across the treatment groups, differences in review times across groups reflect a combination of selection effects (changes in the composition of referees) and behavioral responses (changes in a given referee’s behavior). For instance, referees who expect to be

unable to submit a review quickly might be less likely to agree to review a paper under the shorter four-week deadline. This would reduce average review times in the four-week group via a selection effect even if referee behavior did not change.

Distinguishing between selection and changes in behavior is not critical for a journal editor seeking to reduce average review times, because it does not matter whether improvements come from getting faster referees or inducing a given set of referees to work faster. For the broader objective of learning about how incentives affect prosocial behavior, however, it is important to separate selection from behavioral responses. We therefore begin by assessing selection and then present estimates of treatment effects on review times both with and without adjustments for selection.

We evaluate the magnitude of selection effects in two ways. First, we compare pre-determined referee characteristics, such as tenure status and nationality, across the four groups. We find that these characteristics are generally quite similar across referees who accept invitations in the four groups (details available in Appendix Table B.2).

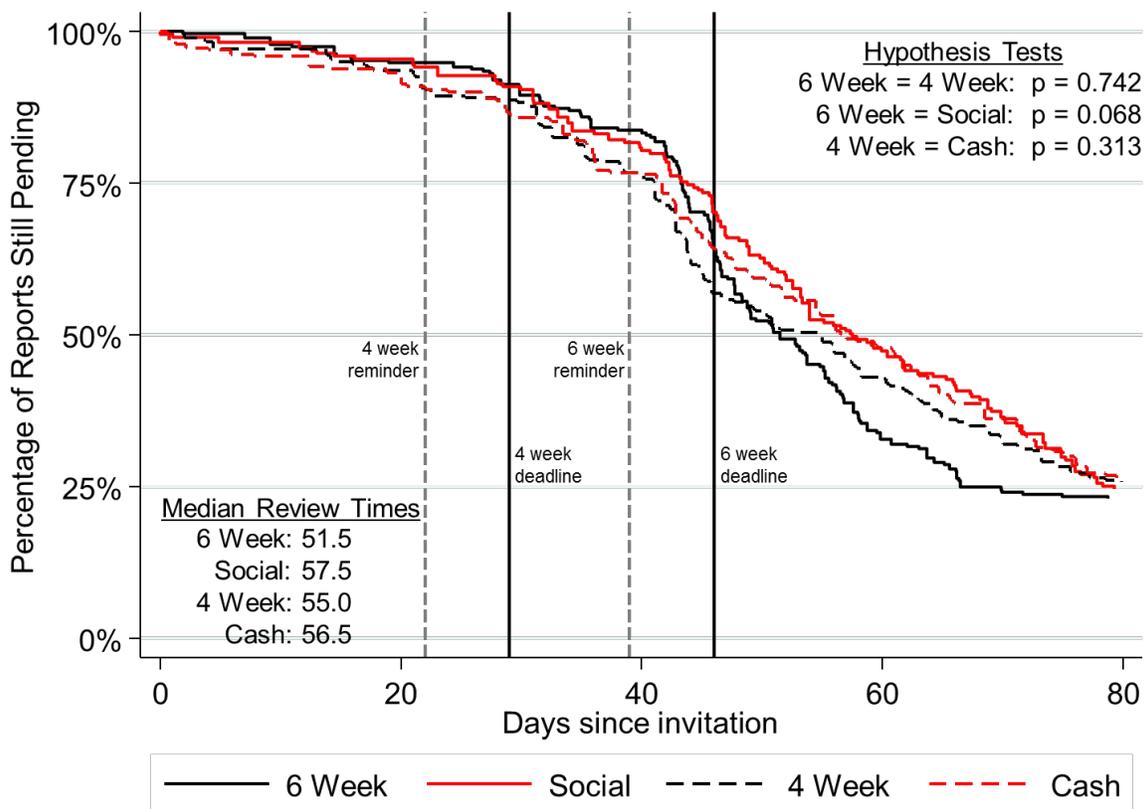
Second, we compare the pre-experiment review times of referees who agreed to review papers in each of the four experimental groups. For this analysis, we focus on the 67 percent of referees in our primary experimental sample who reviewed a manuscript for the journal before the experiment began (from November 2005 to February 15, 2010). All of these pre-experiment reviews were subject to a six week deadline. Figure 2.1 plots survival curves for review times according to the treatment group to which the referees were later assigned, using data from the most recent review before the experiment began. These survival curves show the fraction of reviews that are still pending after a given number of days.<sup>4</sup>

The survival curves in the cash, four-week deadline, and six-week deadline groups are all very similar. Referees who agreed to submit a review under a shorter deadline or cash incentive treatment are no faster than those in the control group based on historical data. Non-parametric (Wilcoxon) tests for equality of the survival curves uncover no differences in

---

<sup>4</sup>We include reviewers who do not submit reviews in these and all subsequent survival curves by censoring their spells at the point when editors make a decision on the paper.

**Figure 2.1:** *Pre-Experiment Review Times for Referees who Accept Invitations During Experiment*



NOTE.—This figure plots survival curves that show the distribution of pre-experiment review times by treatment group. The sample consists of referees who accepted invitations between February 15, 2010 and May 9, 2011, the period when the cash reward was offered. Among these referees, 67.3 percent accepted a review invitation before the experiment began (from November 2005 to February 15, 2010); we use their data to construct this figure. For referees who reviewed multiple papers, we use the most recent pre-experimental review. Each survival curve plots the percentage of reports still pending vs. the number of days elapsed since the referee received the invitation. The solid vertical lines depict the six week deadline (45 days) and the four week deadline (28 days) that were used during the experiment. The dashed vertical lines depict the reminders sent one week before each deadline. Before the experiment, all referees faced the six week deadline and reminders were not sent systematically. We report median review times, defined as the point at which the fraction of reports pending is 50 percent, for each group. We also report p-values from non-parametric Wilcoxon tests for the hypothesis that the pre-experiment review times are the same in each treatment group and its corresponding control group. We compare the four-week and social incentive groups to the six-week group. We compare the cash group to the four-week group because the cash group also faced a four week deadline. We truncate the x-axis at 80 days in the figure for scaling purposes, but use all available data for the hypothesis tests.

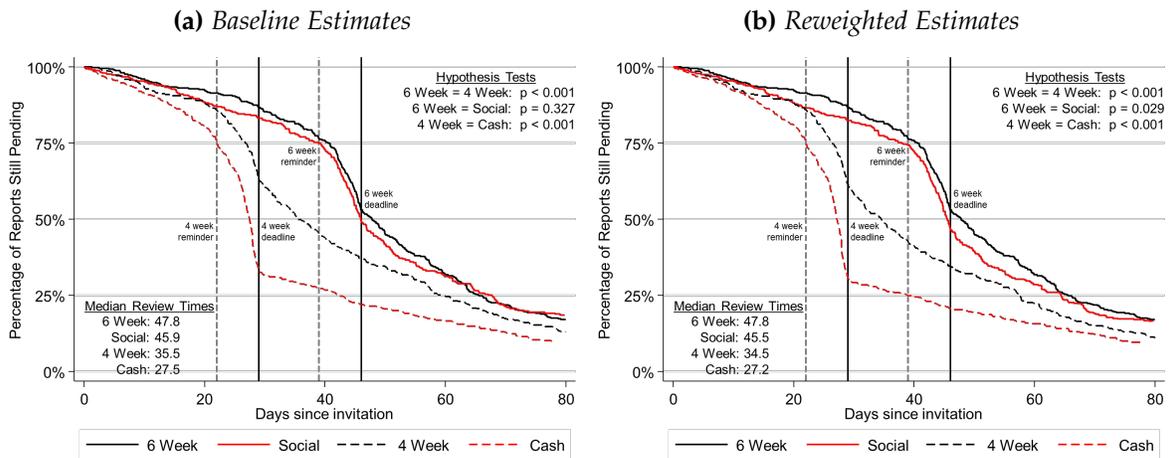
review times across these three groups. We find marginally significant evidence ( $p = 0.068$ ) that referees who agree to review papers in the social incentive group are slightly slower than those in the six-week control group. Hence, if anything, the social incentive treatment appears to induce slightly unfavorable selection in terms of referee speed. One explanation may be that diligent referees tend to be more concerned and anxious about their reputation and are hence less likely to accept the invitation with the social treatment. Overall, this evidence indicates that selection effects are modest and that differences in outcomes across the groups during the experiment are likely to be driven primarily by changes in referee behavior, with the possible exception of the social incentive group.

Figure 2.2 presents our main results on the impact of the treatments on review times during the primary experimental period. Panel 2.2a plots raw survival curves for reviews by treatment group. In Panel 2.2b, we adjust for selection using propensity score reweighting as in DiNardo *et al.* (1996). We reweight the four-week, cash, and social incentive groups to match the six-week group on pre-experiment review times (including an indicator for having no pre-experiment data) using the procedure described in Appendix B.4. We report median survival times (the point at which 50% of reports have been submitted) and non-parametric Wilcoxon tests for the equality of the survival curves in each figure (see Appendix Table B.3 for details).

In contrast with the survival curves in Figure 2.1, the survival curves in Figure 2.2 diverge sharply, showing that the treatments induced substantial changes in review times. Adjusting for differences in prior review times (Panel B) does not affect the results substantially, indicating that most of the change in review times is driven by changes in referee behavior rather than selection effects. We discuss next the impacts of each of the treatments in detail, starting with the shorter deadline and then turning to the cash and social incentives.

Shortening the deadline from six weeks (45 days) to four weeks (28 days) reduces median review times by 12.3 days, based on the baseline estimates in Panel A of Figure 2.2. Hence, we estimate that shortening the deadline by one day reduces median review times by  $12.3 / (45 - 28) = 0.72$  days. The effect is so large because nearly 25 percent of referee reports

**Figure 2.2: Review Times by Treatment Group During Experiment**



NOTE.—This figure plots survival curves showing the distribution of review times by treatment group during the primary experimental period, February 15, 2010 to May 9, 2011 (when the cash reward was offered). In Panel A, each survival curve plots the percentage of reports still pending vs. the number of days elapsed since the referee received the invitation. Panel B replicates Panel A, reweighting the observations in the three treatment groups to match the distribution of pre-experiment review times in the six-week group (see Appendix D for details). The solid vertical lines depict the six week deadline (45 days) and the four week deadline (28 days). The dashed vertical lines depict the reminders sent one week before each deadline. We report median review times, defined as the point at which the fraction of reports pending is 50 percent, for each group. We also report p-values from non-parametric Wilcoxon tests for the hypothesis that review times are the same in each treatment group and its corresponding control group. We compare the four-week and social groups to the six-week group. We compare the cash group to the four-week group because the cash group also faced a four week deadline. We truncate the x-axis at 80 days in the figures, but use all available data for the hypothesis tests.

are submitted in the week between the reminder email and the deadline, and the shorter deadline simply shifts these reports forward. Before week three (shown by the first dashed line in Figure 2.2), the number of pending reports in the four-week and six-week groups is not very different; however, in week four, the survival curve for the four-week deadline group drops sharply relative to the six-week group. The four-week deadline thus appears to act as a nudge that makes referees work on their reports in the fourth week rather than the sixth week.

Providing a \$100 cash incentive for submitting a report within four weeks reduces median review times by an additional eight days relative to the four week deadline. The cash incentive has powerful effects especially after referees receive the reminder email: nearly 50 percent of referees submit a report in the window between the reminder email and the deadline for receiving the cash payment. Missing the four week deadline simply postpones writing the report by a few weeks but costs \$100. Consistent with what one would predict based on a standard model of intertemporal optimization, the survival curve is much flatter immediately after the four week deadline, as very few referees submit reports immediately after the cutoff for the cash payment. Nevertheless, because so many referees make an effort to meet the four week deadline, there are fewer reports pending even 10 weeks after the initial invitation in the cash incentive group relative to all the other groups.

The strong response to the cash incentive in the week before the deadline also supports the view that the cash incentive changes referee behavior, rather than the selection of referees who agree to review, as selection effects would be unlikely to generate such non-linear responses. Indeed, the response to the cash treatment is so large that one can show that selection effects account for very little of the impact using a non-parametric bounding approach, as in Lee (2009). Recall from Table 2.2 that referees in the cash group are  $12.3 = (72.0/64.1 - 1)$  percent more likely to accept review invitations than referees in the four-week group. Assuming that referees who accept the four week invitation would also have accepted the (more attractive) cash invitation, we can bound the selection effect by considering the worst case scenario in which the additional referees who accept the cash

invitation have the shortest spells. For example, 66 percent of referees in the cash group submit their report within 28 days. If we exclude the 12.3 percent fastest referees in the cash group, we obtain a selection-adjusted lower bound of  $(66-12.3)/(100-12.3) = 61$  percent submitting within 28 days. This remains well above the 36 percent of referees who submit a report within 28 days in the four-week group, showing that the difference in review times between the two groups cannot be caused by selection. A similar bounding exercise implies that the difference in review times between the four-week and six-week groups also cannot be due to selection.

Figure 2.2 demonstrates that the direct incentive effect of money outweighs any crowd-out of intrinsic motivation to submit referee reports in a timely manner. To investigate the impact of monetary incentives on intrinsic motivation more directly, we study the behavior of referees for the six months after the cash incentive ended on May 9, 2011. A long literature in social psychology starting with the classic work of Deci (1971) predicts that cash rewards have negative long-run effects on prosocial behavior by eroding intrinsic motivation. Existing evidence for this effect is based primarily on lab experiments (Deci *et al.*, 1999; Frey and Jegen, 2001; Kamenica, 2012). Our experiment offers a new test of this hypothesis in the field that complements earlier work on economic incentives and prosocial behavior in other settings (for example, Gneezy and Rustichini, 2000; Gneezy *et al.*, 2011; Lacetera *et al.*, 2013).

In our application, the prediction from theories in which monetary payments crowd-out intrinsic motivation is that referees who had previously received cash incentives should become slower after they stop receiving cash payments—at least relative to referees in the four-week deadline group, who never received cash payments. We test this hypothesis in Figure 2.3, which plots survival curves for referees assigned to the four-week and cash incentive groups using data before May 9 vs. after May 9, when cash payments ended.<sup>5</sup> The

---

<sup>5</sup>Of the referees who were assigned to the cash incentive group and accepted a review invitation after May 9 (after the cash rewards had ended), 47 percent did not receive an invitation to review a manuscript before May 9. To minimize selection effects, we include these referees in Figure 2.3 even though they never received the cash incentive treatment. The estimates in Figure 2.3 should therefore be interpreted as intent-to-treat estimates. Restricting the sample to the selected subset of referees who received prior invitations yields very

survival curves for the four-week group are similar for invitations before and after May 9, indicating that review times do not vary significantly by invitation date. Referees assigned to the cash incentive group are much less likely to meet the 28 day deadline after May 9 than before May 9, when they were receiving cash rewards. However, there is no evidence that these referees become slower than those in the four-week comparison group, which is what one would expect if intrinsic motivation had been eroded. If anything, it appears that the cash treatment leads to some persistent improvements even after the incentive is removed, perhaps because referees have gotten in the habit of submitting reports slightly sooner.<sup>6</sup> We conclude that the temporary provision of monetary incentives does not have detrimental subsequent effects in the case of peer review.

Next, we turn to the social incentive treatment. We find a significant difference between the social incentive and control group survival curves when reweighting on pre-experiment durations in Figure 2.2b. The difference between the unweighted social and control survival curves in Figure 2.2a is smaller and statistically insignificant. This is because the social incentive treatment appears to induce slightly slower referees to accept review invitations, as shown in Figure 2.1. Once we adjust for this selection effect, we find that the social incentive treatment induces referees to work significantly faster, although the magnitude of the impact remains small. Based on the reweighted survival curves, we estimate that the social incentive reduces the median review time by 2.3 days.<sup>7</sup>

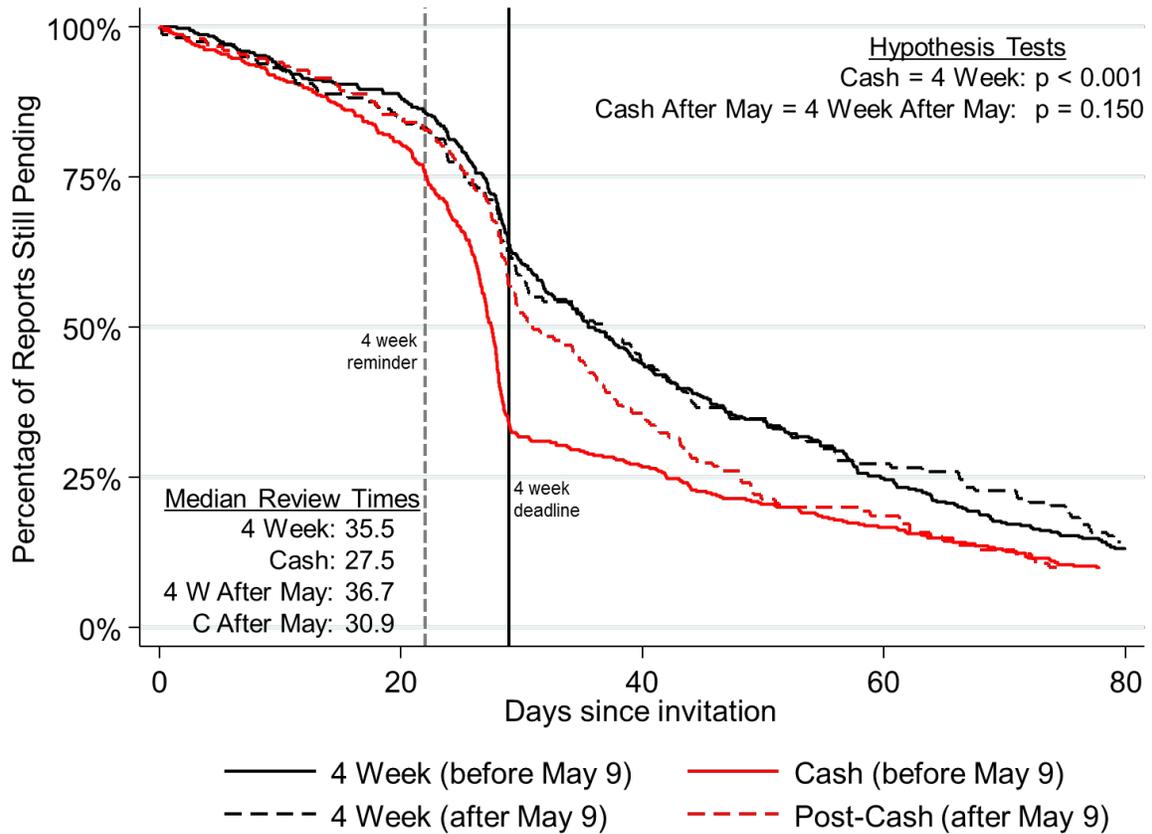
---

similar results.

<sup>6</sup>One might be concerned that referees did not recognize that the cash incentive had stopped after May 9, biasing our comparisons in Figure 2.3. Two facts allay this concern. First, if referees mistakenly thought the cash reward was still in place after May 9, one would expect to see the post-May cash survival curve in Figure 2.3 to drop steeply in the week before the four-week deadline. This does not occur: the post-May cash survival curve tracks the four-survival curves almost perfectly prior to the deadline. Second, the cash incentive increased agreement rates from 64.1 percent (in the four-week group) to 72.0 percent prior to May 9, as shown in Table 2.2. This difference also disappears after May 9: 64.1 percent of referees previously assigned to the cash incentive group agree to do the review after May 9, compared with 65.4 percent in the four-week group during the same period.

<sup>7</sup>We evaluate the robustness of the treatment effect estimates using semi-parametric Cox hazard models in Appendix Table B.4. Consistent with the graphical evidence in Figure 2.2, we find that the cash incentive and 4 week deadlines substantially increase hazard rates of report submission, particularly in the week before the deadline. The social incentive treatment reduces review times significantly when controlling for differences in pre-experiment review times. These results, which are reported in Appendix Table B.4, are robust to changes in

**Figure 2.3:** Review Times Before vs. After End of Cash Reward



NOTE.—This figure plots survival curves showing the distribution of review times in the four-week and cash treatment groups before vs. after May 9, 2011. On May 9, cash rewards were stopped for those in the cash treatment group and referees in this group were subsequently treated identically to those in the four-week group. Hence, the cash (after May 9) group includes referees who previously received cash rewards but no longer do, while the cash (before May 9) group includes referees receiving cash incentives. The four-week group faced the same treatment both before and after May 9. Each survival curve plots the percentage of reports still pending vs. the number of days elapsed since the referee received the invitation. The solid vertical line depicts the four week deadline (28 days). The dashed vertical line depicts the reminder sent one week before the deadline. We report median review times, defined as the point at which the fraction of reports pending is 50 percent, for each group. We also report p-values from non-parametric Wilcoxon tests for the hypothesis that review times are the same in the cash and four week groups before and after May 9.

Finally, we explore the heterogeneity of the treatment effects by referee characteristics. We find no significant heterogeneity in treatment effects by several of the referee characteristics we collected: an indicator for holding an academic position, gender, and an indicator for working in the United States. However, we do find substantial heterogeneity in treatment effects between tenured and untenured referees, as shown in Figure 2.4. This figure replicates Figure 2.2a, dividing the sample into referees who had tenure at the time they were invited to review the manuscript (Panel A) and those who were not tenured at that time (Panel B). The shorter deadline has a significantly larger effect on untenured referees than tenured referees. Untenured referees make a clear effort to submit reports before the deadline, as evident from the sharp drop in the survival curve in Figure 2.4b just before the deadline for the four-week group. In contrast, tenured referees are not very sensitive to the shorter deadline.

The cash incentive improves performance substantially in both groups, but again the impact is larger among untenured referees: 78 percent of untenured referees submit reports before the deadline to receive the cash reward, whereas only 58 percent of tenured referees do so. While the cash incentive and shorter deadline have smaller effects on tenured referees, the social incentive has larger effects on tenured referees. Figure 2.4b shows that review times are almost identical in the social incentive and control groups for untenured referees. In contrast, tenured referees in the social incentive group submit reports significantly earlier than those in the control group, as shown in Figure 2.4a.

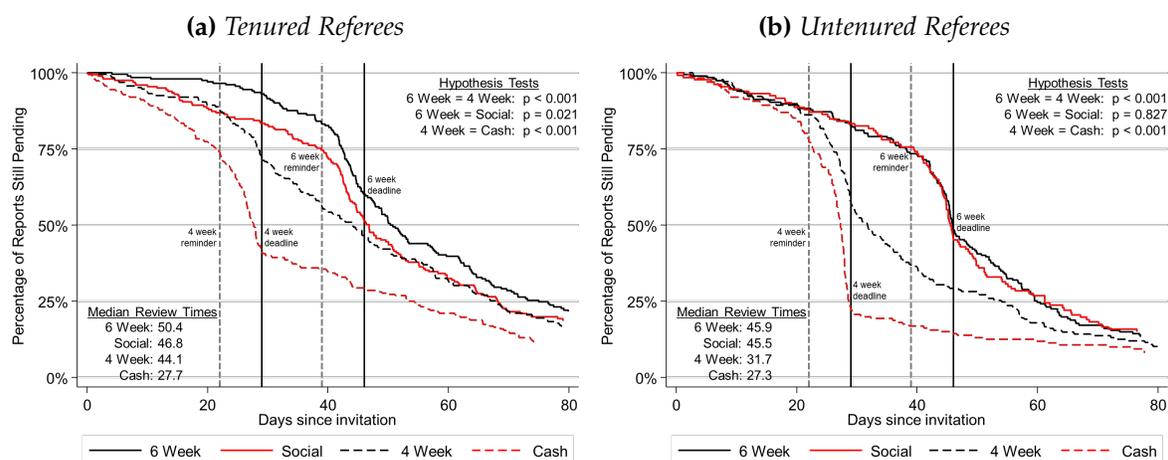
One explanation for why the social incentive treatment is more effective among tenured referees is that untenured referees are already concerned about their reputation with co-editors, who are typically senior colleagues in their field. In contrast, tenured referees might become more concerned about their professional reputation when they face social pressure.<sup>8</sup> Regardless of whether the heterogeneous effects are driven by this mechanism,

---

the control vector and sample specifications. In Appendix Figure 2.2, we use all the data through the end of the experiment (October 26, 2011) rather than restricting the sample to the point at which cash treatments were stopped (May 9, 2011). The point estimates remain similar, but we obtain more precise estimates when using all the data as expected.

<sup>8</sup>Consistent with this explanation, we find that tenured referees are considerably slower than untenured

**Figure 2.4: Heterogeneity in Treatment Effects by Tenure Status**



NOTE.—This figure replicates Figure 2.2a, splitting the sample between tenured referees (Panel A) and untenured referees (Panel B). Tenure status is measured during the experiment based on information from CVs posted online (see Appendix B.3 for details); referees whose tenure status could not be identified are excluded from this figure. In both panels, the sample consists of referees who accepted invitations between February 15, 2010 and May 9, 2011, the period when the cash reward was offered. Each survival curve plots the percentage of reports still pending vs. the number of days elapsed since the referee received the invitation. See notes to Figure 2.2 for further details.

the findings in Figure 2.4 suggest that social incentives can usefully complement other policy instruments by improving behavior among groups who are less responsive to cash incentives and nudges.

### 2.4.1 Outcome 3: Review Quality

Models of multi-tasking predict that if an agent is given an incentive to perform better in one aspect of a job (such as production speed), performance in other aspects of the job (such as quality) might deteriorate. Might the treatments that induce referees to submit reports more quickly also lead referees to submit lower-quality reviews?

We measure the quality of reviews in two ways. The first is an indicator for whether the editor follows the referee’s recommendation with regard to whether the manuscript should be accepted, rejected, or revised and resubmitted. The second is the length of the referee

referees in the control group, but behave like untenured referees in the social incentive group, as shown in Appendix Figure B.3.

report. While length is not equivalent to quality, one natural way in which referees might submit a report more quickly is by providing less detailed comments to authors, especially since only the editor knows the referee's identity.

Table 2.3, which is constructed in the same way as Table 2.2, shows the fraction of cases in which the editor follows the referee's recommendation (Panel A) and the median length of the referee report (Panel B) by treatment group. We find no statistically significant differences across the groups in the rate at which editors follow the referee's advice. We do, however, find that referees write shorter reports to authors under the social and cash incentive treatments. The median report is approximately 100 words (11 percent) shorter in the social and cash groups relative to the six-week and four-week groups. These findings suggest that referees who rush to submit a report earlier because of explicit cash or social incentives might cut back slightly on the level of detail in their comments to authors. Interestingly, referees do not write shorter reports to meet the four-week deadline, consistent with the view that many referees begin writing reports only in the week after they receive a reminder.

Overall, we conclude that one can induce referees to submit reviews more quickly without reducing the quality of reviews significantly. Shorter deadlines have no adverse effect on either measure of quality, while cash and social incentives induce referees to write slightly shorter reports but do not affect the quality of the review as judged by the editor's ultimate decision.

#### **2.4.2 Outcome 4: Spillover Effects on Other Journals**

A natural concern with interventions that improve referee performance at one journal is that they may have negative spillover effects at other journals. Do referees who submit reviews more quickly at the *Journal of Public Economics* prioritize them over other referee reports? In this case, changes in journal policies might not improve the overall efficiency of the review process.

We test for such spillover effects using data from 20 other Elsevier journals in related

**Table 2.3: Measures of Review Quality by Treatment Group**

	6 Week	Social	4 Week	Cash
A. Agreement between Editor Decision and Referee's Recommendation				
Editor Follows Referee's Recommendation	77.9%	76.2%	77.5%	76.2%
	(2.00)	(2.34)	(2.20)	(2.15)
p-value for equality with control		0.585	0.884	0.686
Observations	403	324	373	404
B. Length of Referee Report				
Median Number of Words in Referee Report	877	757	864	786
	(29.1)	(32.5)	(30.3)	(29.2)
p-value for equality with control		0.006	0.757	0.064
Observations	401	321	369	399

NOTE.—This table shows the effects of the treatments on review quality. The sample includes all referees who received invitations sent between February 15, 2010 and May 9, 2011 (the period when the cash reward was offered) and submitted a report. In Panel A, the outcome is the fraction of reports in which the editor's decision (reject vs. accept/revise-and-resubmit) matches the referee's recommendation. We report standard errors in parentheses. Standard errors are clustered by referee in Panel A (but not Panel B). We also report p-values for the null hypothesis that the percentages are the same in each treatment group and its corresponding control group. For the social and 4 week groups, the control group is defined as the 6 week deadline group. For the cash incentive group, the control group is defined as the 4 week deadline group, which is the relevant comparison because the cash incentive group also faced a 4 week deadline. The number of observations (referee reports submitted) is reported in the last row. In Panel B, the outcome is the median number of words in the referee report. Standard errors are reported in parentheses and the p-values are for hypothesis tests analogous to those in Panel A. The number of observations is the number of submitted reports for which we were able to obtain automated word counts of report length.

subfields, such as the *Journal of Health Economics* and the *Journal of Development Economics* (see Appendix B.6 for a complete list). We analyze referee invitations from other journals that are received (1) after referees have received an invitation from the *Journal of Public Economics* during the primary experimental period and (2) before December 31, 2011.

Specifically, we test whether referees' propensities to review manuscripts and their review times at other journals vary across our four treatment groups. Each observation in this analysis is a referee invitation at another journal. The mean agreement rate is approximately 60% in all four groups, with no statistically significant differences across the groups (see Appendix Table B.5). Median review times are approximately 56 days in all four groups, again with no statistically significant differences across the groups (see Appendix Figure B.4).<sup>9</sup>

Of course, referees must postpone some activity to prioritize submitting referee reports. The social welfare impacts of our treatments depend on what activities get postponed. If referees postpone activities with pure private benefits such as leisure, social welfare may increase because referee reports have positive externalities. If on the other hand referees postpone working on their research or on other prosocial tasks, expediting referee reports could reduce welfare. If small delays in these other activities have little social cost, the welfare costs from such delays would be modest. Understanding the nature of crowd-out across different forms of prosocial behavior is an interesting question that we defer to future research.

## 2.5 Lessons for the Peer Review Process

Our results offer three lessons for the design of the peer review process at academic journals.

First, shorter deadlines are extremely effective in improving the speed of the review process. Moreover, shorter deadlines generate little adverse effect on referees' agreement

---

<sup>9</sup>The similarity across the four groups in performance at other journals supports the view that the treatment effects at the *Journal of Public Economics* during the experimental period are driven by changes in referee behavior rather than selection effects.

rates, the quality of referee reports, or performance at other journals. Indeed, based on the results of the experiment, the *Journal of Public Economics* now uses a four week deadline for all referees.

Second, cash incentives can generate significant improvements in review times and also increase referees' willingness to submit reviews.<sup>10</sup> However, it is important to pair cash incentives with reminders shortly before the deadline. Some journals, such as the *American Economic Review*, have been offering cash incentives without providing referees reminders about the incentives; in this situation, sending reminders would improve referee performance at little additional cost.

Third, social incentives can also improve referee performance, especially among subgroups such as tenured professors who are less responsive to deadlines and cash payments. Light social incentives, such as the *Journal of Financial Economics* policy of posting referee times by referee name, have small effects on review times. Stronger forms of social pressure—such as active management by editors during the review process in the form of personalized letters and reminders—could potentially be highly effective in improving efficiency. It would be useful to test this hypothesis in future work using an experiment in which editors are prompted to send personalized reminders to referees at randomly chosen times.

More generally, our findings show that it is possible to substantially improve the efficiency of the peer review process with relatively low-cost interventions, demonstrating the value of studying the peer review process empirically (as in Card and DellaVigna, 2012). Our results reject the view that the review process in economics is much slower than in other fields, such as the natural sciences, purely because economics papers are more complex or difficult to review.

---

<sup>10</sup>These findings contrast with the results of Squazzoni *et al.* (2013), who argue that monetary rewards decrease the quality and efficiency of the review process based on a lab experiment designed to simulate peer review. Our results might differ because the peer review process requires referees to invest considerable time to read papers and write referee reports, unlike the investment game studied in this lab experiment.

## 2.6 Lessons for Increasing Prosocial Behavior

Beyond the peer review process, our results also offer some insights into the determinants of prosocial behavior more broadly.

First, attention matters: reminders and deadlines have significant impacts on behavior. Nudges that bring the behavior of interest to the top of individuals' minds are a low-cost way to increase prosocial behavior, consistent with a large literature in behavioral economics (Thaler and Sunstein, 2008).

Second, monetary incentives can be effective in increasing some forms of prosocial behavior. We find no evidence that intrinsic motivation is crowded out by financial incentives in the case of peer review, mirroring the results of Lacetera *et al.* (2013) in the case of blood donations. While crowd-out of intrinsic motivation could be larger in other settings, these results show that one should not dismiss corrective taxes or subsidies as a policy instrument simply because the behavior one seeks to change has an important prosocial element.

Finally, social incentives can be effective even when other policy instruments are ineffective. This result echoes findings in other settings—such as voting (Gerber *et al.*, 2008), campaign contributions (Perez Truglia and Cruces, 2014), and energy conservation (Allcott, 2011)—, and suggests that social incentives are a useful complement to price incentives and behavioral nudges.

## Chapter 3

# Compensated Discount Functions: An Experiment on the Influence of Expected Income on Time Preference<sup>1</sup>

### 3.1 Introduction

In order to make inferences from their trade-offs between delayed monetary rewards, experimental studies exploring the nature of time preferences typically presume that a subject's marginal utility for money is constant across time. However, several theoretical papers note that subjects may integrate these rewards with their baseline consumption levels (Olson and Bailey, 1981; Rubinstein, 2002; Frederick *et al.*, 2002; Noor, 2009; Gerber and Rohde, 2010). In this case, anticipated changes in marginal utility for money would influence their trade-offs between delayed rewards.<sup>2</sup> This is also related to the recent experimental and theoretical literature that accounts for an unavoidably uncertain future as a contrast to

---

<sup>1</sup>Co-authored with Attila Ambrus, Jawwad Noor, and Tinna Laufey Ásgeirsdóttir

<sup>2</sup>A strand of the literature focuses on timing a consumption stream, or primary rewards (McClure *et al.*, 2007), acknowledging that the value of timed monetary rewards can change with liquidity constraints and other financial frictions, as well as *in vivo* transfers.

a certain present.<sup>3</sup> On the other hand, there is the “narrow bracketing” view that subjects treat experimental rewards in isolation from their background expected financial situation. It may be that integration with background plans is too difficult for people to do,<sup>4</sup> or that small rewards are often viewed as windfalls under different mental accounting and enjoyed separately from base consumption. This important open empirical question of whether background expected financial conditions matter for intertemporal choices is the subject of this paper. Furthermore, we show how to compensate for possible background marginal utility effects when measuring discount functions.

We conducted a lab experiment in Reykjavík, Iceland, using a random sample of individuals from the census, conditional on them living in post codes not too far from the lab.<sup>5</sup> The possibility of prompt access to individual tax information in Iceland allowed us to both observe subjects’ income for the two years preceding and following the experiment.

In the first part of the experiment, we used a standard design of asking participants to choose between unconditional present rewards and unconditional rewards received one or two years later.<sup>6</sup> The elicitation was done through a series of binary choice questions, presented as a standard multiple price list. From this we derive the *uncompensated discount function*: for instance, if a subject is indifferent between \$100 in the current period and \$150 in  $t$  years, we obtain  $D_u(t) = \frac{100}{150}$ . This discount function is uncompensated in the sense that expected income changes can create a wedge between the indifference point and the subject’s true underlying discount function.

---

<sup>3</sup>See Weber and Chapman (2005); Fernandez-Villaverde and Mukherji (2002); Baucells and Heukamp (2012); Andreoni and Sprenger (2012); Halevy (forthcoming).

<sup>4</sup>See some arguments along this line in pages 356-357 of Frederick *et al.* (2002).

<sup>5</sup>This feature of taking lab experiments to field subjects is similar to Andersen *et al.* (2008). Note that such subject pools very likely include more people with stable income than the canonical lab subjects of students. As consumptions of the latter are highly volatile (and steeply rising in a matter of years), any confound of income trends we document in our subsample is likely to be even more powerful for subject pools biased towards students.

<sup>6</sup>We used a front-end delay of one week to put both options on an even footing with respect to transaction costs, immediacy, or trust in the experimenters. What we label here as present rewards were specified to be paid a week after the experiment. Similarly, rewards labeled one and two years later were paid one year plus one week, and two years plus one week later.

In the second part of the design, we asked participants to choose between (a) unconditional present rewards and (b) rewards received one or two years later *conditional* on the subject's income staying "approximately" constant.<sup>7</sup> The idea is that approximately constant income corresponds to approximately constant marginal utility for money, and so the utility evaluation of the rewards is time-independent. Note that the evaluation of a conditional delayed reward depends on the utility from receiving the reward, the degree to which it is discounted due to temporal delay, and the subject's *beliefs* about the likelihood that the conditioning event will be satisfied. Eliciting the second element – the true discount function – is our objective. By assuming that the rewards are small relative to background consumption, we assume approximate linearity of utility from the rewards. We elicit beliefs in an incentivized fashion as follows. First, we sought a good whose utility to the subject is plausibly independent of the marginal utility of money – we consider anonymous charity payments, to a charity of the subject's choice, possess this property. Next, we ask subjects to compare (a') a payment  $h$  made to a charity of the subject's choice at time  $t$  with probability  $\alpha$  and (b') a payment  $h$  to the same charity at time  $t$  conditional on the subject's income staying roughly the same.<sup>8</sup> By observing preferences for various  $\alpha$ , we uncover the desired subjective belief. With this in hand, we can then derive the underlying discount function (see Section 3.2 for details), which we label as the *compensated discount function*, to contrast it with the original uncompensated discount functions that do not take marginal utility expectations into account.

The above design allows us to address the fundamental question of whether expected income levels influence preferences for delayed monetary rewards. With further assumptions, we can make statements about how expected consumption levels (which we do not observe

---

<sup>7</sup>More precisely, we required that the after-tax inflation-adjusted income of the participant stayed within 4% of base income, for both the month and the year preceding the moment in question. See our online appendix for the translation of our actual questionnaire.

<sup>8</sup>As all our experimental questions, these choices were incentivized. At the end of the experiment exactly one binary choice problem was randomly selected for each subject, and the subject's reward was based on the choice in this problem. Hence, some subjects received rewards corresponding to anonymous charity contributions to their charities of choice.

directly) influence these preferences. In particular, our results can be reinterpreted along these lines if we assume that (i) the baseline consumption level is determined by the subject's income, and hence does not change if the subject's income stays the same; (ii) subjects face frictions (these could be external borrowing and lending constraints or internal cognitive constraints) that induce them to consume small rewards at the time when they are received; and (iii) the money amounts we offer subjects are "small" relative to their background consumption, in which case money adds to the agent's background utility approximately linearly. Assumption (ii) is made in most experimental investigations of time preferences. Indeed, if subjects could freely transfer monetary amounts across time periods then they would only care about the discounted present value of their earnings, and appear to the experimenter as exponential discounters with the interest rate as the discount rate. The extensive evidence for non-exponential discounting in the literature is therefore evidence that the consumption of rewards may not be smoothed over time. Regarding assumption (iii), we consider it to be a good approximation, as the maximum reward a subject could earn in our experiment was approximately \$165 (in 2010 currency), which is a small fraction of our typical subject's annual income.

Our first set of results only uses the traditional experimental questions that use unconditional future rewards. Consistent with most existing studies, we find that in the full sample the estimated discount factor between one and two years from the experiment is significantly higher than the estimated discount factor between the time of the experiment and one year later.<sup>9</sup> In the quasi-hyperbolic " $\beta$ - $\delta$ " framework, the estimated present-bias parameter  $\beta$  is 0.89 and significantly different from 1, while the estimated long-run discount factor  $\delta$  is 0.92. We get a very different picture though when we look at subjects with stable incomes. In particular, when we restrict attention to subjects whose real annual income

---

<sup>9</sup>Many experiments describe discount functions as hyperbolic, implying a present bias. See Frederick *et al.* (2002) for a review, and more recent evidence by Benhabib *et al.* (2010). Pender (1996) finds similar results in a field experiment using rice instead of money for rewards. Other papers find no evidence for hyperbolic discounting: see for example Harrison *et al.* (2002); Andersen *et al.* (2008); Andreoni and Sprenger (2012). The experimental finding of present bias generated much theoretical work such as Laibson (1997); O'Donoghue and Rabin (1999); Gul and Pesendorfer (2001); Fudenberg and Levine (2006), and numerous applications built on these models.

remained within a 10% range of current income in both of the two years following the experiment (31 of the 116 subjects), the discount factor between the first and second year after the experiment is almost identical to the discount factor between the present and one year after the experiment, and the estimated  $\beta$  is 0.97 and not significantly different from 1 (while the estimated  $\delta$  for this group is almost exactly the same as for the whole subject population). This means that those subjects whose real income remained stable after the experiment made choices consistent with exponential discounting, while other subjects on average made choices that revealed significant present bias.

We get similar results when, instead of realized income, we use subjects' expectations, elicited in an unincentivized survey, at the time of the experiment to identify those who expect stable income in the two years following the experiment. For example, for those subjects who expect to stay in their current job with more than 80% probability for the two years following the experiment, the estimated discount factor between one and two years from the experiment is not significantly different than the discount factor between the time of the experiment and one year later, and the estimated  $\beta$  is 0.94.

We also investigate subjects whose income increases significantly in the year following the experiment, but stays relatively stable afterwards.<sup>10</sup> Given that the expected marginal utility of extra income for these subjects (as long as they could foresee the said income pattern) is lower in both future years than at the time of the experiment, but about the same magnitude in the two future years, we expect a more significant present bias in these cases. In line with the theoretical predictions, the estimated  $\beta$  decreases to 0.79.<sup>11</sup>

To further examine the relationship between income stability and exponential discounting, in the second half of the analysis we investigated subjects' choices regarding conditional rewards, and derived their compensated discount functions. We encountered two issues

---

<sup>10</sup>In the standard model, a one-shot future reduction (resp. increase) in marginal utility of money predicts present (resp. future) bias in unconditional choices, confounding the true curvature of the utility function with uncertainty and trends in consumption.

<sup>11</sup>There were few subjects with a substantial reduction in income to allow a test of the corresponding prediction of higher estimated  $\beta$ .

needed to be addressed for our data analysis. First, our derivation of the compensated discount functions assumed that subjects followed the discounted expected utility model, and to the extent that this misspecifies some of the subjects' model, we may not get meaningful compensated discount functions for all subjects. Second, the conditional questions are cognitively more demanding than the simpler traditional unconditional questions, and therefore we had to expect that some subjects' responses may not be meaningful (for this purpose we put in much effort to streamline the experimental questionnaire and provide a detailed instruction session). Consequently, we conducted our analysis based not just on the whole sample, but also on a restricted subsample of subjects whose answers conveyed "sensible" compensated discount factors. Without imposing arbitrary restrictions on the data based either on the results or some covariates, we fit a finite mixture of bivariate normal distributions on the one-period discount factors and the ones between years 1 and 2 maximizing the Bayesian Information Criterion, with discrete classification into clusters. The number of mixtures fitted was chosen by the same algorithm and the same criterion, also allowing for classification of observations as outliers (noise). Out of the resulting three clusters, two show compensated discount rates in a reasonable range, consisting of 68.1% of the sample (79 observations overall). The likelihood of falling into either of these two clusters is positively correlated with the subject's education level.

We ran multiple tests of the compensated discount rates being equal to the uncompensated ones (using their difference), or either being 1 (the benchmark of treating future and present equivalently). For  $\beta$  corresponding to present-bias in the quasi-hyperbolic discounting framework, we also tested for differences between the compensated and the uncompensated cases (using their difference), as well as either measure being different from 1. For all these, we employed two-sided  $t$ -tests using standard errors robust to heteroskedasticity, and also non-parametric Wilcoxon signed rank tests.

Our first finding is that compensated discount factors are higher than uncompensated ones, for the entire two-year time period covered in our experiment. This suggests that experiments that do not take into account expectations regarding future income overestimate

subjects' true impatience. Second, we find that compensated discount functions are significantly less present biased than uncompensated ones ( $p = 0.24$  for the arithmetic difference being 0, but  $p = 0.033$  with the signed rank test). The estimated  $\beta$  parameter, when using compensated discount factors, is higher than when using uncompensated discount factors, and not significantly different from 1, while it is significantly less than one when using uncompensated discount factors.

To summarize, our two strands of analysis point in the same direction. Both suggest that people's preferences for monetary rewards are affected by their underlying (expected) income, and that their inherent impatience is less than what traditional experiments eliciting time preferences tend to find. When differences in expected future incomes are compensated for, the implied discount function becomes less hyperbolic, and closer to exponential.

The rest of the paper is structured as follows. Section 3.2 presents the theory of rewards integrated with other income, including our formula for recovering primitives from elicited choices and subjective probabilities. Section 3.3 details the experimental design, Section 3.4 the conduct, Section 3.5 our calculations and empirical strategy. Section 3.6 gives the empirical results, while Section 3.7 concludes.

## 3.2 Theoretical background

Here we provide theoretical foundations for the experiments investigating conditional discount factors.

### 3.2.1 Overview

Suppose that subjects evaluate consumption using an expected discounted utility model with (uncertain) background consumption  $b_t$ , given a differentiable utility function  $u(\cdot)$ . As we compute below, these subjects will evaluate an unconditional reward  $m$  received at time  $t$  by  $D(t) \cdot E[u(b_t + m) - u(b_t)]$ , that is, the discounted increase in expected utility due to the reward. Moreover, assuming that  $m$  is small relative to  $b_t$ , we can approximate the change in

utility as  $u(b_t + m) - u(b_t) = u'(b_t) \cdot m$ . Therefore, the utility of an unconditional reward is

$$D(t) \cdot E(u'(b_t)) \cdot m.$$

Denote current income by  $b^*$ .

Traditional experiments elicit the present amount  $X_t^u$  which is just as good as a future unconditional reward, yielding the (approximate) equality

$$u'(b^*) \cdot X_t^u = D(t) \cdot E(u'(b_t)) \cdot m.$$

Define the *uncompensated discount function* by

$$D_u(t) = \frac{X_t^u}{m}. \quad (3.1)$$

Seeing that  $D_u(t) = D(t) \cdot \frac{E(u'(b_t))}{u'(b^*)}$ , it is clear that the uncompensated discount function correctly estimates the true discount function  $D(t)$  if and only if  $\frac{E(u'(b_t))}{u'(b^*)} = 1$  for all  $t$ , which does not hold for general expectations, and is the reason for the usual assumption of constant marginal utility  $u'(b_t)$  across time in experiments that use  $D_u(t)$  as an estimate for  $D(t)$ .

To compensate for income effects (non-constant marginal utility across time), consider payments that are paid only in the “constant income” event that  $b_\tau = b^*$  for all  $\tau \leq t$ . Denote this event by  $s^t$ . Then such conditional rewards are evaluated by the discounted increase in expected utility:

$$D(t) \cdot p(s^t) \cdot u'(b^*) \cdot m.$$

So, if an agent states that  $X_t^c$  received today is as good as  $m$  received at  $t$  under the condition  $s^t$ , we obtain the equality  $u'(b^*) \cdot X_t^c = D(t) \cdot p(s^t) \cdot u'(b^*) \cdot m$ , and thus,

$$\frac{X_t^c}{m} = D(t) \cdot p(s^t). \quad (3.2)$$

Since  $\frac{X_t^c}{m}$  is observable, we need only elicit  $p(s^t)$  to compute the true discount function  $D$ . A means of deriving such beliefs is to assume the existence of a commodity  $h$  whose utility  $v(h)$  is independent of base consumption  $b_t$  and additively separable from the utility for

money. We consider anonymous charitable contribution to the charity of the agent's choice to be a plausible example of such commodity. Then a charity payment of  $h$  at time  $t$  under the condition  $s^t$  is evaluated by<sup>12</sup>

$$D(t) \cdot p(s^t) \cdot v(h).$$

If the subjects exhibits indifference between such a charity payment  $h$  at time  $t$  under the condition  $s^t$  and an unconditional charity payment  $h$  made at the same time  $t$  but with objective probability  $\alpha^t$ , then it is clear from the implied equality  $D(t) \cdot p(s^t) \cdot v(h) = D(t) \cdot \alpha^t \cdot v(h)$  that

$$\alpha^t = p(s^t).$$

Now the *compensated discount function* can be defined by

$$D_c(t) = \frac{X_t^c}{m} \cdot \frac{1}{\alpha^t}, \quad (3.3)$$

and indeed  $D_c(t) = D(t)$ .

In what follows we describe the model more precisely.

### 3.2.2 Model

Time is discrete and with finite horizon,  $\mathcal{T} = \{0, 1, \dots, T\}$ . The set of possible (inflation-adjusted) future base-consumption levels at any time  $t$  is given by the finite set  $B_t = B \subset \mathbb{R}_+$  with generic element  $b$ . This corresponds to assuming the the set of possible income levels are bounded and measurable in the unit of monetary exchange. Because base-consumption will be uncertain, the set  $B$  is the period  $t$  state space. Period 0 base-consumption is given by  $b^* \in B$  and below we will assume that this is known. The  $t$ -horizon state space is  $\mathcal{S}(t) = \prod_{i=1}^t B_i$ , with generic element  $s^t = (b_0, \dots, b_t)$ . The full state space is  $\mathcal{S} = \cup_{t \in \mathcal{T}} \mathcal{S}(t)$ .

The set of (inflation-adjusted) monetary prizes is an interval  $\mathcal{M} = [0, M]$ . Writing  $B_t = \{b_t^0, \dots, b_t^{N_t}\}$  with  $b_t^0 \leq \dots \leq b_t^{N_t}$ , we let  $M$  be larger than the grid size for base

<sup>12</sup>There is no reason that  $v(h)$  must be discounted by  $D$ , but this is without loss of generality for our purpose.

consumption.<sup>13</sup> In what follows, we will require that  $M$  be small, thereby also requiring a fine state space.

A state-contingent reward is a function  $x : \mathcal{S} \rightarrow \mathcal{M}$  that delivers a prize  $x(s^t) \in \mathcal{M}$  at date  $t$  conditional on the realization of  $s^t = (b_0, \dots, b_t)$ . The set of all state-contingent rewards is denoted  $X$ . The primitive of our analysis is a preference  $\succsim$  on  $X$ .

We maintain relatively weak assumptions in this section just to show that our experimental procedure is not tied to a particular functional form for discounting. But the reader should keep in mind that all the experimental analysis in this paper will be based on the well-known beta-delta model.

*Basic assumptions* – The subject is assumed to evaluate future consumption according to a discounted utility model where uncertainty is evaluated according to (state-dependent) subjective expected utility theory. Instantaneous utility is given by a strictly increasing and differentiable function  $u : \mathbb{R}_+ \rightarrow \mathbb{R}$  with a differentiable inverse. The discount function is  $D(t) > 0$  and satisfies  $D(0) = 1$  but is not necessarily strictly monotone or restricted to take values less than 1.<sup>14</sup> The subject's prior (over future base consumption) is a probability measure  $p$  on  $\mathcal{S}(T)$ .

*Integration assumptions* – Assume that the subject integrates state-contingent rewards with her anticipated base consumption and completely consumes any prize in the period and state that it is received. The presumption here is that the rewards in  $\mathcal{M}$  are small enough for this to be an acceptable assumption. It follows that the discounted expected utility due to a state-contingent reward  $x$  given beliefs  $p$  is

$$U(x) = u(b^*) + \sum_{(b_1, \dots, b_T) \in \mathcal{S}(T)} \left[ \sum_{t=0}^T D(t) u(b_t + x(b_0, \dots, b_t)) \right] p(b_0, \dots, b_T). \quad (3.4)$$

*Full support assumption* – We assume that  $p(s^T) > 0$  for all  $s^T$ , that is, unconditional beliefs on each  $B_t$  have full support, and positive probability is assigned to future base

---

<sup>13</sup>That is,  $M > \max\{|b_t^i - b_t^{i+1}| : t \leq T, i < N_t\}$ .

<sup>14</sup>If  $D(t) > 1$  is permitted then in order to ensure the existence of present equivalents we assume that  $u$  is unbounded.

consumption staying the same as current consumption  $b^*$ . Given the strict monotonicity of  $u$ , this is equivalent to assuming behaviorally that for any  $s_T$ , there exists some prize  $m > 0$  such that the state contingent reward  $x_{m,s^T}$  that pays  $m$  at  $T$  in state  $s^T$  and 0 otherwise satisfies

$$x_{m,s^T} \succ \omega,$$

where  $\omega$  denotes the state-contingent rewards that yields 0 at all  $t$  and  $s^t$ .

### 3.2.3 Deriving $D$ from $\succsim$

The primitive preference  $\succsim$  on  $X$  has the representation (3.4) where the utility of a reward  $x(s^t)$  received at time  $t$  conditional on  $s^t$  is *state-dependent*.<sup>15</sup> That is, a dollar received in a given period depends on base consumption  $b$  in that period, so that the instantaneous utility in that period is  $u(b+1)$ , and thus dependent on  $b$ . This might lead one to suspect that the representation (3.4) lacks desirable uniqueness properties, and in particular the key component of interest, the discount function  $D$ , may not be unique. This would be problematic since it would imply that  $D$  is not pinned down by preferences  $\succsim$ , and in particular, there is no meaningful sense in which it can be extracted from  $\succsim$  in any experiment.<sup>16</sup> Therefore, we must establish that any discounted expected utility representation for  $\succsim$  must have a unique  $D$ . This is the content of the proposition below, the proof of which is relegated to the appendix.

**Proposition 1.** *The prior  $p$  and the discount function  $D$  are uniquely determined by  $\succsim$ .*

---

<sup>15</sup>Though the preference over consumption streams has a state-independent representation, the induced representation for the preference over state-contingent rewards is state-dependent.

<sup>16</sup>For instance, in state-dependent subjective expected utility, a function  $f$  that takes states into prizes is evaluated by  $\sum_s u(f(s), s)p(s)$ . In this representation, the prior is not unique and thus has no behavioral meaning. We could take any  $\alpha_s > 0$  for each  $s$ , take a monotone transformation of  $u(\cdot, s)$  given by  $v(\cdot, s) = \frac{1}{\alpha_s} u(\cdot, s)$  for all  $s$ , and adopt a different prior given by  $q(s) = \frac{\alpha_s}{\sum_s \alpha_s p(s)} p(s)$  for all  $s$ . Then it is easy to see that the utility function  $f \mapsto \sum_s v(f(s), s)q(s)$  represents precisely the same preference as before. In contrast, when  $u$  is state-independent (as in Savage's subjective expected utility theory), *every* subjective expected utility representation for the preference must share the same prior  $p$ , and thus  $p$  is uniquely pinned down by preferences. It can be elicited by asking the agent to choose between bets.

Having established the possibility of eliciting  $p$  and  $D$  from  $\succsim$  we now outline a procedure for doing so.

Let  $x_{m,s^t}$  be the reward that yields prize  $m$  at time  $t$  conditional on constant base consumption  $s^t = (b^*, \dots, b^*)$ . Denote by  $\psi(x_{m,s^t})$  the reward that yields a prize immediately such that  $\psi(x_{m,s^t}) \sim x_{m,s^t}$ .<sup>17</sup> Identify the immediate prize with  $\psi(x_{m,s^t})$ . The representation (3.4) implies that

$$u(b^* + \psi(x_{m,s^t})) - u(b^*) = D(t)p_t(s^t)[u(b^* + m) - u(b^*)].$$

Note that since  $u$  is a strictly increasing differentiable function with a differentiable inverse,  $\psi(x_{m,s^t})$  is a strictly increasing differentiable function of  $m$  that takes the value 0 when  $m = 0$ . Taking a derivative of the above expression with respect to  $m$  yields

$$u'(b^* + \psi(x_{m,s^t})) \frac{\partial \psi(x_{m,s^t})}{\partial m} = D(t)p_t(s^t)u'(b^* + m).$$

Evaluating at  $m = 0$  gives  $\frac{\partial \psi(x_{m,s^t})}{\partial m} \Big|_{m=0} = D(t)p_t(s^t) \frac{u'(b^*)}{u'(b^*)}$  and so

$$\frac{\partial \psi(x_{m,s^t})}{\partial m} \Big|_{m=0} = D(t)p_t(s^t).$$

In practice we can rely on an approximation via the observation that for small  $m$ ,

$$\frac{\partial \psi(x_{m',s^t})}{\partial m} \Big|_{m'=0} \approx \frac{\psi(x_{m,s^t}) - \psi(x_{0,s^t})}{m - 0} = \frac{\psi(x_{m,s^t})}{m}.$$

Hence

$$\psi(x_{m,s^t}) \approx D(t)p_t(s^t)m$$

Note that  $\psi(x_{m,s^t})$  and  $m$  are observable, so if we can identify  $p_t(s^t)$  (for instance, as described earlier) then we can find  $D(t)$ .

---

<sup>17</sup>The existence of such a reward is implied by the unboundedness and continuity of  $u$ . Its uniqueness is implied by the strict monotonicity of  $u$ .

### 3.2.4 Discussion: Interpretations and Assumptions

The theoretical considerations above suggest an experimental design, that will be presented in detail in the next section. To be able to interpret the experiment as measuring how expected background consumption affects time preferences, and as a valid way of uncovering the underlying time preferences, one needs to make the following assumptions.

1. Rewards are consumed fully in the period of receipt.
2. The marginal utility of rewards is linear for the range of rewards that we consider.
3. In case income remains unchanged, background consumption does not change.

The first assumption is a standard assumption in the traditional time preference experimental literature. It conflicts with the consumption smoothing property implied by the standard life cycle model, a testable implication of which is that all rewards should be ranked according to its present value and elicited discount functions must be exponential (with a discount rate equal to the market rate of interest). Indeed some studies find no evidence of hyperbolic discounting for money while simultaneously confirming hyperbolic discounting for consumption (Augenblick *et al.*, 2013). But the fact that we, like many studies, reject nonexponential discounting it follows that perfect smoothing is not going on.<sup>18</sup> Zero smoothing of small rewards then serves a standard benchmark, which can be justified in several ways.<sup>19</sup>

In our experiment a period is *one year*, and thus the rewards we consider (the maximal possible reward was approximately \$165 in 2010 currency) are very small compared with annual consumption. This speaks to our second assumption, which is standard in the traditional time preference literature, but see Andersen *et al.* (2008) for a recent critique.

The third assumption is necessitated by the fact that consumption is difficult to observe directly. Consistent with this assumption, empirical studies show that consumption tracks

---

<sup>18</sup>The heterogeneity of results across experiments could potentially be due to income effects that are not accounted for.

<sup>19</sup>There may be transactions costs or cognitive costs of revising lifetime consumption. Moreover, there may exist a temptation for a splurge that one's financial situation does not justify, and an optimal response to this may well be to fully consume small windfalls (for instance see Fudenberg and Levine (2006)). Indeed, for such reasons a sophisticated agent may consume small windfalls.

income over the life cycle (Browning and Crossley, 2001)), pointing both to possible frictions in consumption smoothing and to possibly boundedly rational behavior.

We note that even if the assumptions relating with consumption do not hold, the experimental design is still valid in testing whether expected future income influences preferences for delayed monetary rewards. Such dependence would drive a wedge between our elicited compensated and uncompensated discount functions.

### 3.3 Experimental Design

The experiment used a questionnaire with three sections, all three translated fully in the online appendix. In the first one we asked subjects to choose between rewards received at different times. Some of the future rewards involved payments that were conditional, i.e. only received if the subject's real income remains "approximately constant" (defined shortly) up to the time of payment. The second part provided subjects binary choice questions involving a payment to the subject's charity of choice either with an exogenously given probability or under the condition the subject's real income remains "approximately constant". Finally, the third part featured an unincentivized questionnaire eliciting demographic and financial information about subjects, as well as subjective beliefs about future income. All in all, this amounted to six sets of decisions that all included a range of binary choices before they turned to a survey – the small number of questions helps keeps the cognitive burden low on the subjects. During instructions at the beginning of the experiment, we explained to the subjects how a dice roll would determine which one of those six choice categories will be used to generate payoffs, and for that choice one random line from the range of binary choices, according to their expressed preferences.<sup>20</sup>

The first section started with questions involving only unconditional payments. A generic question asks for the subject's preference between<sup>21</sup>

---

<sup>20</sup>It is a dominant strategy to make the binary choices truthfully, with the caveat that if a subject is exactly indifferent between the two options then the choice can be either of them.

<sup>21</sup>In 2010, 20,000 ISK was worth approximately \$165.

(i) a “later” payment of 20,000 Icelandic Kronur (ISK), paid to the subject at  $t$  years plus 1 week later,

(ii) a “present” payment in the amount of  $x$  paid 1 week later,

where  $t=1$  or 2 years, and  $x$  ranges from ISK 200 to ISK 22,000 in steps of ISK 200. This series of binary choice questions is presented as a standard multiple price list, that is, in the form of a table.<sup>22</sup> Although subjects could indicate their preference in each cell of the table, for their convenience they were allowed a shortcut where they could indicate two consecutive cells on the table where preferences switch from favoring the present reward to preferring the later reward. Consequently, the subject’s indifference point was captured within an ISK 200 interval.

In order to elicit compensated discount functions (and specifically to bring equation (3.2) into play), we next asked analogous binary choice questions where some payments are paid on the condition that the subject’s income remains “approximately constant”. We formally defined this condition to consist of two requirements:

(a) the price-indexed disposable annual income of the subject during the year following the experiment (in case of a 2-year delayed reward, in both years) is within 4% of the annual income in the 12 months preceding the experiment;

(b) the price-indexed disposable monthly income of the subject in the last month before payment occurs is within 4% of the monthly income in the month preceding the experiment. In the case of payments two years from the experiment, this has to hold true both one year after the experiment and two years after the experiment.

The idea behind part (a) is that the general income level of the subject remains the same, relative to the time of the experiment, while the motivation for part (b) is to make sure that the subject’s overall financial situation is similar to the time of the experiment. We refer to both conditions holding together as the subject’s income situation remaining constant.

---

<sup>22</sup>This is equivalent to a Becker-DeGroot-Marschak (BDM) procedure. Like most of the related literature, we opted for the list of binary questions because we believe the original BDM procedure (in particular, understanding why truth-telling is weakly dominant) to be cognitively more demanding for the subjects. This is also in accordance with our experience from several pilots conducted before the experiment.

Given this definition, the next set of binary choice questions asked subjects to indicate their preference between:

(i') a "conditional later" payment of 20,000 Icelandic Kronur (ISK) paid to the subject, after  $t$  years plus 1 week, *if* income remains approximately constant.

(ii') a "present" payment in the amount of  $x$  paid 1 week later,

where as before,  $t=1$  or 2 years, and  $x$  ranges from ISK 200 to ISK 22,000 in steps of ISK 200.

Section II obtained the data that allows us to exploit equation (3.3). The section started with the subject choosing a charity from a list of well known and established charities with different objectives that were briefly described to the subjects. Subjects were told that the forthcoming questions involve rewards in the form of payments to their charity of choice. They were then asked to indicate their preference between:

(i'') a "conditional charity payment" of ISK 20,000 to the subject's charity of choice, after time  $t$  years plus 1 week, *if* income remains "approximately constant",

(ii'') a "random charity payment" of ISK 20,000 to the subject's charity of choice, after time  $t$  years plus 1 week, with exogenous probability  $p$ ,

where  $t=1$  or 2 years, and  $p$  ranges from 0.01 to 1 in steps of 0.01. This series of questions was presented in the form of a table as in Section I.

In Section III, subjects were asked to fill out a survey, which asked them, among other things, for their bank information in order to transfer their payments. Using the one week delay in payments, as opposed to exactly at the time of the experiment, as well as one and two years later, allowed us to use the exact same procedures and payment methods, regardless of whether the rewards were delayed or not. This section also included questions on the subjects' social and economic background, as well as expectations on their future economic situation. It also included a direct but unincentivized question on how likely they think their yearly income remains approximately the same one and two years following the experiment. We also elicited subjects' probability assessments of entering a new job by one and two year's time after the experiment, and their probability assessments on losing their current jobs.

## 3.4 Experimental Procedures and Background

### 3.4.1 Experimental Sessions

The experiment took place on June 9th and 10th of 2010. Recruitment was conducted by phone from a random sample of Icelanders between the ages of 20 and 45 living in western or central Reykjavík, specifically post codes 101, 105 and 107. The sample was collected from the census by Skýrr, an Icelandic IT company and frequent government contractor. Subsequently, the subjects' phone numbers were collected manually through ja.is, the online Icelandic Telephone Directory. The experiment was conducted in groups simultaneously in one location, specifically a lecture hall at the University of Iceland. The hall had a podium and an overhead projector used in the presentation of instructions, which was carried out by one of the researchers (photos of the location are available upon request). Before starting, the subjects were asked to read and sign a consent form. They were also asked not to talk to each other and informed that if they had questions they should raise their hand, rather than speak up, and they would be assisted individually by a researcher or an assistant. Each session consisted of approximately 15 subjects and took a little bit over one hour. Outside the classroom we set up four dice rolling stations at which assistants reported the randomized outcome of the subjects' dice roll and computerized randomization process.

Before the experimental session in 2010, we conducted several small (10-25 subjects) informal pilots in 2007-2009. These pilots featured similar questions as the experiment, but subjects only received a fixed compensation for participating, independent of their answers. These pilots were mainly used to fine-tune how to effectively explain the questions in the experiment.<sup>23</sup>

---

<sup>23</sup>We also conducted two small post-experiment pilots, one in 2011 and one in 2012, with slightly altered questionnaires and instructions as the ones used in the experimental session in 2010, to investigate whether these design changes lead subjects to a better understanding of the questions involving conditional rewards. As we did not find any evidence for this, these post-experimental pilots did not lead to a subsequent incentivized experimental session.

### 3.4.2 Payment Process

As online banking is widespread in Iceland, subjects were paid by bank transfer. In all instances, both for pilots and the experiment itself, subjects received their payments. This happened in the vast majority of instances at the scheduled time. In a few instances with illegible account numbers, payments made with a few days' delay after quick follow-up e-mails or phone calls. Payments were initiated by the finance division at the University of Iceland. In 2010, 48 subjects in the experiment received payments, on average ISK 14,823. In 2011, 17 payments were made, in the amount of ISK 20,673. 12 of those went directly to participants, but five went to charities of their choice. In 2012, 13 individuals received a payment of ISK 21,891 each<sup>24</sup>.

### 3.4.3 Income Verification

We chose Iceland for the experiment because of the prompt availability of comprehensive income-tax information due to a pay-as-you-earn system where the income tax is continuously withheld at source. That is, the lion's share of income-tax revenue in Iceland is collected monthly, and the Directorate of Internal Revenue (DIR) receives fairly accurate accounts of each individual's income in a timely fashion. For this reason, we signed a contract with the DIR on February 12th 2010, in accordance with the Icelandic Data Protection Act 77/2000 and a notification to the Icelandic Data Protection Authority (S4052). According to this DIR contract and the subjects' informed consent, we did not obtain direct information on subjects' incomes. Instead, income *changes* were calculated by the DIR staff, and they sent us the percentage changes in subjects' monthly as well as yearly incomes, for the specific months and twelve month periods. This was done using the latest information available in the DIR systems, which is generally fairly complete by the 18th of the following month, with only minor adjustments after that.

According to our contract, DIR calculated and delivered the income changes at three

---

<sup>24</sup>These delayed payments were inflation-adjusted equivalents of ISK 20,000 at the time of the experiment.

points in time. During the week after the experiment took place, as well as one and two years later.<sup>25</sup> Income changes were adjusted for changes in the Consumer Price Index (CPI) which is published by Statistics Iceland by the second to last day of the reference month.

#### **3.4.4 Economic Situation Around the Time of the Experiment**

The seemingly flourishing economy of Iceland suffered a major meltdown less than two years before our main experimental session, when the country's three largest banks collapsed and were nationalized. In a widely-viewed televised address, Prime Minister Geir Haarde announced to the country: "(T)here is a very real danger, fellow citizens, that the Icelandic economy, in the worst case, could be sucked with the banks into the whirlpool and the result could be national bankruptcy" (Prime Minister's Office, 2008). Although a sovereign default did not follow, this is indicative of the volatility and uncertainty of the situation. During the following months, hundreds of firms in the country declared bankruptcy. The announcement of the crisis triggered international consequences, including a decision by the United Kingdom to freeze the assets of one of the three large banks (Landsbanki), emergency funding from the International Monetary Fund, protests and a subsequent fall of the government in February 2009.

This was a dramatic macroeconomic shock that affected the entire population of this small open economy with its own currency and for which exchange rates and prices changed suddenly and dramatically. Although the experiment and the post-experimental payments took place after the collapse, and our design in principle remains valid no matter what expectations subjects have regarding future income, the crisis should be noted as many subjects may have felt considerable uncertainty as to how the economy would adjust in the coming years. Iceland is one of the world's smallest currency areas, making the Icelandic krona very vulnerable, which affects the price level and real wages. From the time of the experiment and to the payment dates one and two years later, the price level rose by 4.2%

---

<sup>25</sup>We alerted DIR of each deadline with two week's advance notice, and delivery and payments went through with no delay.

and 9.9% respectively and the real wage rate increased by 2.7% and 4.1% respectively. All in all, it can thus be said that most subjects in our sample faced considerable uncertainty at the time of the experiment regarding their future income.

Our means of deriving compensated discount functions is immune to the economic turbulence as long as subjects use discounted expected utility with respect to some beliefs about the future. To the extent that the ambiguity about the future led subjects to assess the future uncertainty and beliefs in an inconsistent way, it would be revealed to us through unreasonable compensated discount functions.

### 3.5 Computation of Discount Factors and Statistical Analyses

Recall our notation  $X_t^u$  and  $X_t^c$  in Section 3.2 for the present payments that are indifferent to a later conditional and unconditional payment at  $t$  respectively. For simplicity, we estimate these as the lowest payment that the subject indicates as superior to the later payment. For the questions involving only unconditional rewards, if the later payment was paid  $t$  years after the experiment, this gives us an uncompensated discount function  $D_u(t)$  as defined in equation (3.1). It will be convenient to define uncompensated discount factors between  $t$  and  $t + 1$ :

$$D_u(t, t + 1) = \frac{D_u(t + 1)}{D_u(t)}.$$

Note that  $D_u(0) = 1$  so  $D_u(0, 1) = D_u(1)$ . When interpreting these parameters in the standard  $\beta - \delta$  framework, we set  $\delta_u = D_u(1, 2)$ , and

$$\beta_u = \frac{(D_u(1))^2}{D_u(2)} = \frac{D_u(1)}{D_u(1, 2)}.$$

To compute compensated discount functions  $D_c(t)$ , defined by equation (3.3), we first define  $p_t$  as the probability that a subject expects her income  $t$  years after the experiment to be approximately the same as at the time of the experiment (as defined in Section 3.2 and in the experimental instructions). Specifically, this is coded as the lowest reported probability with which a probabilistic contribution to her charity of choice  $t$  years after

the experiment is indicated as superior to a contribution of the same amount at the same time to the same charity, conditional on her income staying approximately the same. The probabilities computed this way, from incentivized experimental questions, strongly correlate with subjects' reported probabilities of stable income in Section III of the experimental questionnaire, although the latter, unincentivized measures are noisier. Linear regressions of the unincentivized responses on the incentivized ones yield slopes 0.602 and 0.442 with an  $R^2$  of 0.3368 and 0.2154, for the two time horizons respectively.

Given  $p_t$ , and  $X_t^c$ , which denotes the amount of present payment that makes the subject indifferent to a conditional payment of ISK 20,000  $t$  years after the experiment, the estimated compensated discount function is  $D_c(t) = \frac{X_t^c}{p_t \cdot 20,000}$ . As with uncompensated discount functions, we define the discount factor  $D_c(t, t+1) = \frac{D_c(t+1)}{D_c(t)}$ ,  $\delta_c = D_c(1, 2)$ , and  $\beta_c = \frac{(D_c(1))^2}{D_c(2)}$ .

For the possible reasons outlined in the introduction – namely the possibility of model misspecification and the possibility of imperfect responses by subjects to the more involved conditional questions – the compensated discount functions computed for a subset of our subjects were not plausible. For example, while it is plausible that a subject may exhibit negative discount rates (that is,  $D_c(t) > 1$ ), it is implausible that discount functions may be steeply upward sloped, as we observed in some cases. Indeed, several instances of implausible discount functions were typically the result of hard-to-rationalize choices such as preferring a reward  $m$  at time  $t$  with probability  $\alpha < 1$  to a reward  $m$  at  $t$  for sure.

In light of this, we proceed by following two separate strategies to analyze our predictions for uncompensated and compensated answers. First, for the simpler, uncompensated choices, we report discount functions by strata of income changes, with different predictions under our assumptions that baseline income levels matter for time preferences over monetary rewards. Besides the full sample, those subsamples are (a) people who experienced relatively stable income for two years after the experiment and would thus be expected to show less present bias under the proposed theory, (b) those who have stable income for two years before and after the experiment and are thus assumed to be individuals with even greater

stability of income and thus even less confound in the conventional measures of present bias, (c) those who, at the time of the experiment, assess their probability of having a new job to be small and should thus show a smaller present bias than the full sample, and finally (d) those individuals whose realized income rises in the first year after the experiment took place but plateaus after that, who should, according to theory show greater present bias than the full sample or any of the subsamples described above.

Second, for the compensated questions, we identify groups with plausible conditional discount functions, and repeat our analysis restricted to these subsamples. Not to force any arbitrary judgment (and thus results) on the data, we employ the tools of statistical cluster analysis to identify (latent) classes of subjects in the data and focus on groups with a “reasonable” range of discount factors. We choose to define clustering in terms of the discount factors  $D_c(0,1)$  and  $D_c(1,2)$  (the latter only implied by  $D_c(1)$  and  $D_c(2)$ ).

For the purposes of clustering, we assume jointly normally distributed discount factors, within an unobserved  $k$  class of subjects. Allowing for noise (outliers), the likelihood function that is numerically maximized is

$$\prod_{i=1}^n \left[ \frac{\tau_0}{V} + \sum_{k=1}^G \tau_k \phi_k(\mathbf{x}_i | \theta_k) \right], \quad (3.5)$$

where  $V$  denotes the hypervolume of the data region, data can come from different distributions or simply be noise, which have respective probabilities  $\tau_k$  (and thus of course  $\tau_k \geq 0$  and  $\sum_k \tau_k = 1$ ). The algorithm initializes with noise estimates coming from a nearest-neighbor method and hierarchical clustering applied to the rest of the data (with a simple maximization using the EM algorithm), and the EM algorithm alternating Bayesian updating conditional on the parameter estimates (Expectation step) and maximizing in the parameters conditional on the classification probabilities (Maximization step).

In equation 3.5 the likelihood contribution for an observation comes from assuming that the densities of the discounts factors follow a bivariate normal distribution. This procedure is conditional on the number of clusters,  $G$ . This we let be chosen to maximize the Bayesian

Information Criterion (BIC),

$$BIC = 2 \cdot \log L(\mathbf{x}, \theta^*) - (\#\text{parameters}) \log n. \quad (3.6)$$

*Formal hypotheses* We conduct a family of tests over two important variants of two measures, using a parametric and a nonparametric test, both two-sided and also one-sided about the economically interesting differences.

Our first set of tests compare first-year and second-year discount factors. The tests are done for both compensated and uncompensated, so we drop the superscript on  $D$  below to simplify exposition.

- $H_0 : D(1,2) = D(0,1)$ , corresponding to subjects being exponential discounters.
- $H_1 : D(1,2) > D(0,1)$ , corresponding to subjects exhibiting present bias in their time preferences.

We also report test results of no present bias without taking the stance on which way a deviation could occur. This amounts to a two-sided test against the alternative hypothesis of inequality.

Our preferred parametric test is a paired  $t$ -test, testing the null hypothesis that  $H_0 : E[D(1,2)] = E[D(0,1)]$  against the alternative  $H_1 : E[D(1,2)] > E[D(0,1)]$  or  $H_1 : E[D(1,2)] \neq E[D(0,1)]$ . The test makes the usual assumption that the discount factors are normally distributed in small samples, or the samples are large enough that the asymptotic approximation is good enough. The nonparametric test of our choice is a sign test, testing the medians without assuming the two variables have the same distribution: comparing the null hypothesis that  $H_0 : \text{med}[D(1,2)] = \text{med}[D(0,1)]$  against the alternative  $H_1 : \text{med}[D(1,2)] > \text{med}[D(0,1)]$  or  $H_1 : \text{med}[D(1,2)] \neq \text{med}[D(0,1)]$ .

We conduct the tests both for the uncompensated and the compensated discount factors.

Finally, we repeat the exact same test procedures for another parametrization of present bias, namely the transformation of the discount factors into the common  $\beta = \frac{D(0,1)}{D(1,2)}$  pa-

parameter being larger or less than one.<sup>26</sup> Thus for the  $t$ -tests the null is  $H_0 : E[\beta] = 1$  (no present bias on average) against the alternatives of  $H_1 : E[\beta] < 1$  or  $H_1 : E[\beta] \neq 1$ , while for the sign tests, the analogues with medians. Again, we conduct the tests separately for uncompensated or compensated measures.

## 3.6 Experimental Results

### 3.6.1 Uncompensated Discount Factors

The bottom pane of Table 3.1 shows the difference of first and second year uncompensated discount factors ( $D_u(1,2) - D_u(0,1)$ ), as well as the implied  $\beta$  and  $\delta$  parameters, with the strata by income stability around the time of the experiment and times of payment in separate columns. There are 31 subjects in column 2, whose realized annual real income in both years after the experiment stayed within 10% of real income at the time of the experiment. Not surprisingly, these are slightly older subjects, more likely to be employed, but also better educated and there are more females than in the full sample (The top pane of Table 3.1 shows summary statistics using our survey in Section III). Reassuringly, these subjects were much less likely to expect job changes after the experiment.<sup>27</sup> Column 3 focuses on subjects whose real income was stable both before and after the experiment. Because most of our subjects experienced turbulent income in the years before the experiment, due to the financial crisis and the subsequent recession, we define this category as the subjects whose real income in both years before the experiment was within 20% of real income at the time of the experiment, and whose real income in both of the two years after the experiment was within 10% of real income at the time of the experiment. Even this more permissive

---

<sup>26</sup>Of course, for the parametric test the transformation can matter because of the approximation being better or worse in our finite samples. The nonparametric test is indifferent to such a transformation.

<sup>27</sup>For those in Column 2 whose realized income stayed stable for two years after the experiment, the assessment (at the time of the experiment) of being in a new job one year (two years) after the experiment is 20.0 (25.6) percentage points less likely than those not having such stable income, and this difference has a  $p$ -value of 0.003 (0.00). For those in column 3 who we categorized as having stable income both before and after the experiment, these differences from the rest of the sample are 26.8 and 27.1 in the same direction (with both  $p$ -values of 0.00).

criterion for the two years before the experiment results in only a small number of subjects (10, with 9 females) being in this category. They are also better educated and slightly older.

We assume that those whose realized income was stable for two years after the experiment expected more stable income in these years than the rest of the subjects. This is consistent with subjects' reported beliefs in Section III of the questionnaire, and the highly correlated incentivized responses in Section II. Yet if we define stable income subjects directly based on their reported expectations in the questionnaire, instead of their realized income that we directly observe, we get results similar to what follows, but noisier (Column 4 focuses on the 32 people who estimated the risk of have a new job in two years to be less than 20%). This suggests that at least some subjects have trouble reporting their expectations in terms of probabilities.

As shown in the bottom pane of Table 3.1, the difference between the second and first year discount factors is large and highly statistically significant for the entire sample of 115 (we lose one observation with an invalid second-period discount factor). The overall sample shows considerable present-bias (average  $\beta$  is 0.895), and a reasonable amount of impatience (on average  $\delta$  being 0.922). In contrast to this, for those 31 people whose incomes were stable for two years after the experiment, the difference in discount factors is statistically insignificant, with a point estimate close to 0, and the implied  $\beta$  is only insignificantly below 1, with a point estimate of 0.972. On the other hand, the estimated  $\delta$  (0.924) is essentially the same as for the whole subject pool. We see the same pattern for those subjects whose income was stable both before and after the experiment.

This difference of mean present-bias parameter  $\beta$  between those with stable income after the experiment and the rest of the subjects is statistically significant at the 5% level ( $p = 0.031$ ), and robust to controlling for other differences between this group and others. Applying the tests in a regression framework, the mean difference in  $\beta$  declines only slightly, from 0.108 to 0.1, when demographic controls from Table 3.1 are included, to control for selection on observables.<sup>28</sup> However, this exercise is informative only if any remaining

---

<sup>28</sup>For this we also include those 8 observations who have BMI, employment or number or children data

**Table 3.1: Description of the Sample**

	Full sample	Stable income after 2010 <sup>†</sup>	Stable income before and after 2010 <sup>††</sup>	New job lesslikely than 20% <sup>†††</sup>	Income rises, then plateaus <sup>††††</sup>
Summary Statistics (with SDs)					
Female	0.59 (0.49)	0.68 (0.48)	0.90 (0.32)	0.59 (0.50)	0.64 (0.50)
Age	33.19 (6.90)	36.68 (6.37)	37.20 (5.35)	35.66 (6.99)	32.18 (7.36)
Higher education	0.43 (0.50)	0.55 (0.51)	0.70 (0.48)	0.44 (0.50)	0.45 (0.52)
Smoker	0.20 (0.40)	0.23 (0.43)	0.20 (0.42)	0.13 (0.34)	0.27 (0.47)
BMI	25.43 (4.52)	24.55 (3.48)	23.93 (2.77)	25.34 (4.27)	25.35 (4.27)
Employed	0.88 (0.33)	0.94 (0.25)	0.90 (0.32)	0.88 (0.34)	1.00 (0.00)
Single	0.60 (0.49)	0.74 (0.44)	0.80 (0.42)	0.72 (0.46)	0.45 (0.52)
Number of children	1.22 (1.36)	1.35 (1.36)	1.30 (1.42)	1.78 (1.48)	1.45 (1.57)
Expect job change in t=1	35.44 (35.46)	20.73 (28.74)	11.00 (15.78)	2.78 (3.96)	20.09 (24.72)
Expect job change in t=2	47.76 (36.64)	28.93 (30.55)	23.00 (26.37)	4.91 (5.40)	42.73 (36.36)
Time Preference Variables (with SEs)					
$D_u(1,2) - D_u(0,1)^*$	0.13 (0.03)	0.04 (0.04)	0.03 (0.06)	0.08 (0.04)	0.22 (0.10)
$\beta_u^{**}$	0.89 (0.03)	0.97 (0.04)	0.98 (0.06)	0.94 (0.04)	0.79 (0.10)
$\delta_u^{***}$	0.92 (0.02)	0.92 (0.02)	0.92 (0.03)	0.90 (0.03)	0.89 (0.06)
N	116	31	10	32	11

NOTE.—The table reports means and standard deviations (in parentheses) for basic descriptive variables for various subsamples in the top pane, means and standard errors (in parentheses) for time preference variables in the bottom pane. \* Reports the average difference between each subject's revealed uncompensated discount factor over the second year and the one over the first year. \*\* Reports the average ratio of discount factors over the second year and over the first year. \*\*\* Reports the average ratio of the two-year discount factor and the one-year discount factor. <sup>†</sup> Only including those whose realized annual income remained within 10% of their 2010 income both in 2011 and 2012. <sup>††</sup> Only including those whose annual incomes in 2008 and 2009 as well. <sup>†††</sup> Includes those whose self-report the probability of getting a new job in either year after the experiment to be less than 20%. <sup>††††</sup> Includes those whose realized annual income rose more than 10% in 2011 but less than 10% on top of that in 2012.

selection on unobservables has an effect comparable to that of the observables. In the framework of Oster (2014) for linear regressions, we can calculate that the improvement of the model fit  $R^2$  by 5% when controlling with observables implies that the group with stable income has the larger  $\beta$  unless (similarly confounding) unobservables could improve  $R^2$  by 50% or more. Such a huge increase in explanatory power is not very plausible, thus even this more cautious perspective leads us to the conclusion that incomes stable after the experiment are associated with less present bias.<sup>29</sup>

We can contrast these results with the 11 subjects who experienced a relatively large real income rise (more than 10%) in the first year after the experiment, but then saw their incomes stabilized (remained within 10% of the income in the first year after the experiment).<sup>30</sup> If these subjects foresaw this income path,<sup>31</sup> their expected marginal utilities for monetary rewards in both of the two years following the experiment are lower than at the time of the experiment. This should imply more present bias than for the rest of the subject pool (without compensating for the confound of the income process, of course). We find some evidence for this, as the estimated  $D_u(1,2) - D_u(0,1)$  difference rises to 0.223, and the estimated  $\beta$  decreases to 0.792 for this subsample, although we have no power in this small sample to establish these changes as statistically significant.

---

missing. These observations enter with value zero for the respective variable, but also with an extra indicator in the regression to take out their group-specific mean separately. An alternative model without these controls have less explanatory power but yield a stable-unstable difference only larger than the unconditional point estimate.

<sup>29</sup>All analogous calculations are collected in Appendix Table C.1. This table contrasts measures of time-preference for these groups in univariate and multivariate regressions to show how robust these differences are. With explanatory power taken into account, the table also reports the other extreme bound on the contrast from Oster (2014), as well as the minimal explanatory power a model with unobservables would need to have to call the sign of the unconditional point estimate confounded. These calculations assume that unobservables correlate with the group indicator “in exactly the same direction” as observables ( $\delta = 1$  in that framework, not to be confused with the time-preference parameter). The bounds on the contrast come from assuming that the all-inclusive models could achieve the theoretical maximum  $R^2 = 1$ .

<sup>30</sup>Without a one-time change in marginal utility of money the model does not make any relevant predictions for uncompensated discount functions.

<sup>31</sup>This is reasonable in many instances, e.g. for those finishing school or for some other reason expecting to get into a higher-paying job.

### 3.6.2 Compensated Discount Factors

#### Clustering outcomes

The noise and clusters resulting from the procedure in Section 3.5 are summarized in the top pane of Table 3.2. The best-fitting distribution is one with no correlations between the two discount rates in any cluster, and three clusters with different means and variances for the normal distributions, and some outliers (noise).<sup>32</sup> Descriptive statistics for clusters are shown in the middle pane of Table 3.2. The first two clusters, containing 79 of the 116 subjects, have compensated discount factors mostly in a plausible range. Since Cluster 1 still contains some individuals with high compensated discount factors for some period of time, we also focus on cluster 2 separately, although this group only contains 16% of our sample. The “reasonable” clusters are slightly better educated, on average (see Table 3.2), which is consistent with the hypothesis that these subjects understood the questions involving conditional rewards better, but the overlap between all clusters is apparent for any covariate.

#### Main results

Comparisons of uncompensated and compensated discount factors are reported in the bottom pane of Table 3.2, with standard errors in parentheses. The p-values of our tests outlined above are reported in Table 3.3. As with the average discount factors before, some statistics for the whole dataset are quantitatively implausible, although they imply the same qualitative conclusions as we obtain when restricting attention to subjects in the two reasonable clusters. We find the same clear patterns no matter whether we include all subjects, only those in Clusters 1 and 2, or only those in Cluster 2. Compensated discount factors, on all horizons are significantly higher than uncompensated ones:  $D_c(0,1) > D_u(0,1)$  and  $D_c(1,2) > D_u(1,2)$ . This suggests that standard elicitation of time preferences, which do

---

<sup>32</sup>The conclusions are robust to different initializations of the clustering. Though the number of clusters might be higher with a different initialization (e.g. initializing the expected outliers differently), the clusters with “reasonable” discount factors largely overlap, as do the vector of noise-indicators.

**Table 3.2: Parameters by Clusters**

	Full sample	Cluster 1	Cluster 2	Clusters 1&2	Cluster 3	Noise
Estimated Parameters of Clustering						
Probability*		0.497	0.159		0.198	0.147
$E[D_c(0,1)]^{**}$		1.374	1.008		1.692	
$V[D_c(0,1)]^{***}$		0.670	0.001		2.702	
$E[D_c(1,2)]^{**}$		1.150	0.971		4.107	
$V[D_c(1,2)]^{***}$		0.183	0.005		9.123	
Summary Means (SDs)						
Female	0.59 (0.49)	0.63 (0.49)	0.53 (0.51)	0.61 (0.49)	0.52 (0.51)	0.56 (0.51)
Age	33.19 (6.90)	33.48 (7.29)	31.84 (6.27)	33.09 (7.05)	32.62 (7.02)	34.44 (6.23)
Higher education	0.43 (0.50)	0.47 (0.50)	0.53 (0.51)	0.48 (0.50)	0.43 (0.51)	0.19 (0.40)
Smoker	0.20 (0.40)	0.23 (0.43)	0.16 (0.37)	0.22 (0.41)	0.14 (0.36)	0.19 (0.40)
BMI	25.43 (4.52)	25.22 (4.55)	25.32 (5.85)	25.24 (4.87)	25.20 (2.98)	26.67 (4.59)
Employed	0.88 (0.33)	0.88 (0.33)	0.84 (0.37)	0.87 (0.34)	0.81 (0.40)	1.00 (0.00)
Single	0.60 (0.49)	0.62 (0.49)	0.74 (0.45)	0.65 (0.48)	0.52 (0.51)	0.50 (0.52)
Number of children	1.22 (1.36)	1.25 (1.36)	1.11 (1.15)	1.22 (1.31)	1.00 (1.34)	1.50 (1.63)
Expect job change	35.44	27.52	33.68	29.04	50.62	46.31

in t=1

**Table 3.2:** (Continued)

	(35.46)	(30.31)	(38.87)	(32.47)	(38.19)	(39.10)
Expect job change	47.76	39.76	43.42	40.66	65.24	59.00
in t=2	(36.64)	(34.63)	(44.07)	(36.90)	(30.39)	(34.09)
Sample Estimates (SEs)						
$D_c(0,1)$	5.963 (1.666)	1.399 (0.108)	1.005 (0.007)	1.305 (0.084)	1.767 (0.408)	34.471 (9.539)
$D_c(1,2)$	2.832 (0.530)	1.145 (0.056)	0.975 (0.016)	1.104 (0.043)	4.510 (0.682)	9.161 (3.296)
$D_c(0,1) - D_u(0,1)$	5.171 (1.664)	0.578 (0.103)	0.226 (0.046)	0.493 (0.081)	1.037 (0.396)	33.692 (9.519)
$D_c(1,2) - D_u(1,2)$	1.926 (0.536)	0.237 (0.056)	0.092 (0.037)	0.202 (0.044)	3.541 (0.672)	8.212 (3.323)
$D_c(1,2) - D_c(0,1)$	-3.131 (1.803)	-0.254 (0.123)	-0.030 (0.018)	-0.200 (0.094)	2.743 (0.992)	-25.310 (11.838)
$D_u(1,2) - D_u(0,1)$	0.130 (0.027)	0.088 (0.031)	0.104 (0.053)	0.092 (0.027)	0.240 (0.084)	0.171 (0.080)
$\beta_u$	0.895 (0.026)	0.929 (0.031)	0.899 (0.056)	0.922 (0.027)	0.791 (0.079)	0.898 (0.082)
$\beta_c$	16.075 (8.997)	1.487 (0.175)	1.037 (0.020)	1.379 (0.135)	1.634 (0.714)	107.595 (62.053)
N	116	60	19	79	21	16

NOTE.—The table collects the parameters corresponding to the clusters as well as means and standard errors (in parentheses) for key constructs. \*Estimated probabilities of the cluster in the mixture model. \*\*Estimated mean of the compensated discount factors. \*\*\* Estimated variance of the discount factors.

not take into account future income expectations of subjects, on average overestimate the amount of impatience of individuals. One explanation for this is that most subjects expect a

rising income path, decreasing the marginal utility of small monetary rewards in the future, relative to the present.

Second, the left pane of Table 3.3 shows that our tests reject exponential discounting without compensation for income changes in the entire sample, but not for the groups with stable income, restating the conclusions of our previous analysis reported in Table 3.3. More interestingly, the new results on compensated discount factors in the right pane of Table 3.3 show no evidence of present bias once we control for the potential confounds of income changes. This suggests that once one compensates for income expectations, there is insufficient evidence that the average subject has a present bias.

### **3.7 Conclusion**

We find that in a setting where most people reveal present bias, people with stable incomes are standard exponential discounters. Our results show that subjects' choices over delayed rewards depend on their income expectations, revealing their importance for intertemporal choice experiments. Our experimental design shows how choices on rewards conditional on no income changes, alongside incentivized elicitations of subjective income expectations, can help researchers compensate for this confound. Our results cast doubt on the maintained hypothesis of many other studies that mental accounting implies that subjects evaluate monetary rewards independently of other income.

It is important for future research to obtain a richer empirical picture on how time preferences over monetary payments depend on different characteristics of future income expectations such as trends in expected income, trends in volatility of income, or the amount of autocorrelation of income in future periods. Any conclusion on saving behavior or policy would be overly speculative at this point, but for fundamental work in choice theory, this is an important step towards more conclusive lab experiments and empirical studies.

**Table 3.3:** *p-values from Hypothesis Tests*

Uncompensated Discount Factors					
Null hypothesis	Test	Full sample	Stable income after 2010 <sup>†</sup>	Stable income before and after 2010 <sup>††</sup>	Income rises, then plateaus <sup>†††</sup>
$D_u(1,2) \leq D_u(0,1)$	t-test	0.000	0.161	0.318	0.023
	sign test	0.032	0.828	0.945	0.113
$\beta_u \geq 1$	t-test	0.000	0.242	0.368	0.034
	sign test	0.032	0.828	0.945	0.113
$D_u(1,2) - D_u(0,1) = 0$	t-test	0.000	0.322	0.637	0.046
	sign test	0.064	0.572	0.344	0.227
$\beta_u = 1$	t-test	0.000	0.483	0.736	0.067
	sign test	0.064	0.572	0.344	0.227
N		115	31	10	11
Compensated Discount Factors					
		Full sample	Cluster 1	Cluster 2	Clusters 1&2
$D_c(1,2) \leq D_c(0,1)$	t-test	0.957	0.979	0.948	0.982
	sign test	0.719	0.817	0.945	0.906
$\beta_c \leq 1$	t-test	0.952	0.996	0.958	0.997
	sign test	0.719	0.817	0.945	0.906
$D_c(1,2) - D_c(0,1) = 0$	t-test	0.085	0.043	0.105	0.036
	sign test	0.699	0.519	0.344	0.282
$\beta_c = 1$	t-test	0.097	0.007	0.083	0.006
	sign test	0.699	0.519	0.344	0.282
N		115	60	19	79

NOTE.—The table collects p-values for tests with the null hypothesis corresponding to present bias. Low p-values thus indicate hyperbolicity or simply “non-exponentiality”. <sup>†</sup> Only including those whose realized annual income remained within 10% of their 2010 income both in 2011 and 2012. <sup>††</sup> Only including those whose annual incomes in 2008 and 2009 as well. <sup>†††</sup> Includes those whose realized annual income rose more than 10% in 2011 but less than 10% on top of that in 2012.

# References

- ALESSIE, R., DEVEREUX, M. P. and WEBER, G. (1997). Intertemporal consumption, durables, and liquidity constraints: A cohort analysis. *European Economic Review*, **41** (1), 37–59.
- ALLCOTT, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, **95** (9-10), 1082–1095.
- ANDERSEN, S., HARRISON, G. W., LAU, M. I. and RUTSTRÖM, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, **76** (3), 583–618.
- and NIELSEN, K. M. (2011). Participation constraints in the stock market: Evidence from unexpected inheritance due to sudden death. *Review of Financial Studies*, **24** (5), 1667–1697.
- ANDREONI, J. and SPRENGER, C. (2012). Risk preferences are not time preferences. *American Economic Review*, **102** (7), 3357–3376.
- AUGENBLICK, N., NIEDERLE, M. and SPRENGER, C. (2013). *Working Over Time: Dynamic Inconsistency in Real Effort Tasks*. Working paper, National Bureau of Economic Research.
- BAUCELLS, M. and HEUKAMP, F. H. (2012). Probability and time trade-off. *Management Science*, **58** (4), 831–842.
- BÉNABOU, R. and TIROLE, J. (2006). Incentives and prosocial behavior. *American Economic Review*, **96** (5), 1652–1678.
- BENHABIB, J., BISIN, A. and SCHOTTER, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior*, **69** (2), 205–223.
- BERTRAND, M., DUFLO, E. and MULLAINATHAN, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, **119** (1), 249–275.
- BROWNING, M. and CROSSLEY, T. F. (2001). The life-cycle model of consumption and saving. *Journal of Economic Perspectives*, **15** (3), 3–22.
- and — (2009). Shocks, stocks, and socks: Smoothing consumption over a temporary income loss. *Journal of the European Economic Association*, **7** (2), 1169–1192.
- CAMPBELL, J. Y. (2013). Mortgage market design. *Review of Finance*, **17** (1), 1–33.
- CARD, D. and DELLA VIGNA, S. (2012). *Revealed Preferences for Journals: Evidence from Page Limits*. Working Paper 18663, National Bureau of Economic Research.

- CARROLL, C. D. (1997). Buffer-stock saving and the life cycle/permanent income hypothesis. *Quarterly Journal of Economics*, **112** (1), 1–55.
- (2009). Precautionary saving and the marginal propensity to consume out of permanent income. *Journal of Monetary Economics*, **56** (6), 780–790.
- CHETTY, R. (2008). Moral hazard versus liquidity and optimal unemployment insurance. *Journal of Political Economy*, **116** (2), 173–234.
- and FINKELSTEIN, A. (2013). Social insurance: Connecting theory to data. In A. J. Auerbach, R. Chetty, M. Feldstein and E. Saez (eds.), *Handbook of Public Economics*, vol. 5, Amsterdam; New York: Elsevier, pp. 111–193.
- and SZEIDL, A. (2007). Consumption commitments and risk preferences. *Quarterly Journal of Economics*, **122** (2), 831–877.
- and — (2010). *The Effect of Housing on Portfolio Choice*. Working Paper 15998, National Bureau of Economic Research.
- CROSSLEY, T. F. and LOW, H. (2011). Borrowing constraints, the cost of precautionary saving and unemployment insurance. *International Tax and Public Finance*, **18** (6), 658–687.
- DAVIDOFF, T. (2010). Home equity commitment, and long-term care insurance demand. *Journal of Public Economics*, **94** (1), 44–49.
- DEATON, A. (1991). Saving and liquidity constraints. *Econometrica*, **59** (5), 1221–1248.
- DECI, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, **18** (1), 105–115.
- , KOESTNER, R. and RYAN, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, **125** (6), 627–668.
- DiNARDO, J., FORTIN, N. M. and LEMIEUX, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, **64** (5), 1001–1044.
- EINAV, L., FINKELSTEIN, A., RYAN, S. P., SCHRIMPF, P. and CULLEN, M. R. (2013). Selection on moral hazard in health insurance. *American Economic Review*, **103** (1), 178–219.
- EJRNAES, M. and HOCHGUERTEL, S. (2011). *Entrepreneurial Moral Hazard in Income Insurance*. Working paper, Tinbergen Institute.
- ELLISON, G. (2002). The slowdown of the economics publishing process. *Journal of Political Economy*, **110** (5), 947–993.
- ENGEN, E. M. and GRUBER, J. (2001). Unemployment insurance and precautionary saving. *Journal of Monetary Economics*, **47** (3), 545–579.
- FEHR, E. and FISCHBACHER, U. (2003). The nature of human altruism. *Nature*, **425** (6960), 785–791.

- FELDSTEIN, M. and ALTMAN, D. (2007). Unemployment insurance savings accounts. In J. M. Poterba (ed.), *Tax Policy and the Economy*, vol. 21, Cambridge, MA: MIT Press, pp. 35–64.
- FERNANDEZ-VILLAVARDE, J. and MUKHERJI, A. (2002). *Can We Really Observe Hyperbolic Discounting?* Working Paper 02-008, University of Pennsylvania.
- FREDERICK, S., LOEWENSTEIN, G. and O'DONOGHUE, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, **40** (2), 351–401.
- FREY, B. S. and JEGEN, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, **15** (5), 589–611.
- FUDENBERG, D. and LEVINE, D. K. (2006). A dual-self model of impulse control. *American Economic Review*, **96** (5), 1449–1476.
- GERBER, A. and ROHDE, K. I. M. (2010). Risk and preference reversals in intertemporal choice. *Journal of Economic Behavior and Organization*, **76** (3), 654–668.
- GERBER, A. S., GREEN, D. P. and LARIMER, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, **102** (1), 33–48.
- GNEEZY, U., MEIER, S. and REY-BIEL, P. (2011). When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, **25** (4), 191–209.
- and RUSTICHINI, A. (2000). A fine is a price. *Journal of Legal Studies*, **29** (1), 1–17.
- GROSS, D. B. and SOULELES, N. S. (2002). Do liquidity constraints and interest rates matter for consumer behavior? evidence from credit card data. *Quarterly Journal of Economics*, **117** (1), 149–185.
- GRUBER, J. (1997). The consumption smoothing benefits of unemployment insurance. *American Economic Review*, **87** (1), 192–205.
- GUL, F. and PESENDORFER, W. (2001). Temptation and self-control. *Econometrica*, **69** (6), 1403–1435.
- HALEVY, Y. (forthcoming). Time consistency: Stationarity and time invariance. *Econometrica*.
- HANSEN, G. D. and IMROHOROUÇLU, A. (1992). The role of unemployment insurance in an economy with liquidity constraints and moral hazard. *Journal of Political Economy*, **100** (1), 118–142.
- HARRISON, G. W., LAU, M. I. and WILLIAMS, M. B. (2002). Estimating individual discount rates in denmark: A field experiment. *American Economic Review*, **92** (5), 1606–1617.
- HENDREN, N. (2013). Private information and insurance rejections. *Econometrica*, **81** (5), 1713–1762.
- HOLMSTROM, B. and MILGROM, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, **7**, 24–52.

- HURST, E. and STAFFORD, F. (2004). Home is where the equity is: Mortgage refinancing and household consumption. *Journal of Money, Credit, and Banking*, **36** (6), 985–1014.
- JENSEN, T. L. R., REN LETH-PETERSEN, S. and NANDA, R. (2014). *Housing Collateral, Credit Constraints, and Entrepreneurship-Evidence from a Mortgage Reform*. Working Paper 20583, National Bureau of Economic Research.
- KAMENICA, E. (2012). Behavioral economics and psychology of incentives. *Annual Review of Economics*, **4** (1), 427–452.
- LACETERA, N., MACIS, M. and SLONIM, R. (2013). Economic rewards to motivate blood donations. *Science*, **340** (6135), 927–928.
- LAIBSON, D. I. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, **112** (2), 443–477.
- LEDYARD, J. (1995). Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth (eds.), *Handbook of Experimental Economics*, vol. 1, Princeton: Princeton University Press, pp. 111–194.
- LEE, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, **76** (3), 1071–1102.
- LENTZ, R. (2009). Optimal unemployment insurance in an estimated job search model with savings. *Review of Economic Dynamics*, **12** (1), 37–57.
- LETH-PETERSEN, S. (2010). Intertemporal consumption and credit constraints: Does total expenditure respond to an exogenous shock to credit? *American Economic Review*, **100** (3), 1080–1103.
- MCCLURE, S. M., ERICSON, K. M., LAIBSON, D. I., LOEWENSTEIN, G. and COHEN, J. D. (2007). Time discounting for primary rewards. *Journal of Neuroscience*, **27** (21), 5796–5804.
- NOOR, J. (2009). Hyperbolic discounting and the standard model: Eliciting discount functions. *Journal of Economic Theory*, **144** (5), 2077–2083.
- O'DONOGHUE, T. and RABIN, M. (1999). Doing it now or later. *American Economic Review*, **89** (1), 103–124.
- OLSON, M. and BAILEY, M. J. (1981). Positive time preference. *Journal of Political Economy*, **89** (1), 1–25.
- OSTER, E. (2014). *Unobservable Selection and Coefficient Stability: Theory and Evidence*. Working paper, University of Chicago Booth School of Business.
- PARSONS, D. O., s, T. T. and r, H. B. L. (2003). *Voluntary Public Unemployment Insurance*. EPRU Working Paper Series 03-05, Economic Policy Research Unit (EPRU), University of Copenhagen. Department of Economics.
- PENDER, J. L. (1996). Discount rates and credit markets: Theory and evidence from rural india. *Journal of Development Economics*, **50** (2), 257–296.

- PEREZ TRUGLIA, R. and CRUCES, G. (2014). *Social Incentives in Contributions: Field Experiment Evidence from the 2012 U.S. Presidential Campaigns*. Working paper, Microsoft Research.
- RUBINSTEIN, A. (2002). *Comments on the Risk and Time Preferences in Economics*. Working Paper 6, Tel Aviv University Foerder Institute for Economic Research Sackler Institute for Economic Studies.
- SHAPIRO, M. D. and SLEMROD, J. (2003). Consumer response to tax rebates. *American Economic Review*, **93** (1), 381–396.
- SHIMER, R. and WERNING, I. (2008). Liquidity and insurance for the unemployed. *American Economic Review*, **98** (5), 1922–1942.
- SQUAZZONI, F., BRAVO, G. and TAKÁCS, K. (2013). Does incentive provision increase the quality of peer review? an experimental study. *Research Policy*, **42** (1), 287–294.
- THALER, R. H. and SUNSTEIN, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven: Yale University Press.
- TITMUSS, R. M. (1971). *The Gift Relationship: From Human Blood to Social Policy*. London: George Allen and Unwin.
- VESTERLUND, L. (2012). *Voluntary Giving to Public Goods: Moving Beyond the Linear VCM*. Working paper, Harvard University.
- WEBER, B. J. and CHAPMAN, G. B. (2005). The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? does delay eliminate the certainty effect? *Organizational Behavior and Human Decision Processes*, **96** (2), 104–118.

# **Appendix A**

## **Appendix to Chapter 1**

### **A.1 Supplementary Tables**

**Table A.1:** *Sample Selection*

Selection criteria	Number of Individuals
Cohorts 1957-1962	452,583
Homeowner in 1992	202,561
Drop if moved in 1992 before reform	195,418
Trim financial outliers	189,021
Balance sample	183,251
Drop if out of labor force	162,139
Drop if self-employed	144,619
Drop if in education	133,785
Drop if insufficient observations to calculate unemployment risk	133,779
Drop if missing industry code	119,952
Drop if missing labor market experience	118,018
Trim top & bottom 1% of liquidity shock	115,656
Trim top & bottom 1% of permanent income	113,344

NOTE.—The table describes the number of observations retained in each step of our sample restrictions, as described in the main text.

**Table A.2:** Cohorts Affected by Early Retirement Reforms of the Unemployment Insurance System

Cohort	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	Note
1920	67	68	69	70	71	72	73	74	75	76	They do not need insurance anyway – they can just retire at will if they lose their job
1921	66	67	68	69	70	71	72	73	74	75	
1922	65	66	67	68	69	70	71	72	73	74	
1923	64	65	66	67	68	69	70	71	72	73	
1924	63	64	65	66	67	68	69	70	71	72	
1925	62	63	64	65	66	67	68	69	70	71	
1926	61	62	63	64	65	66	67	68	69	70	
1927	60	61	62	63	64	65	66	67	68	69	
1928	59	60	61	62	63	64	65	66	67	68	
1929	58	59	60	61	62	63	64	65	66	67	
1930	57	58	59	60	61	62	63	64	65	66	
1931	56	57	58	59	60	61	62	63	64	65	
1932	55	56	57	58	59	60	61	62	63	64	
1933	54	55	56	57	58	59	60	61	62	63	Transitional allowance into ER
1934	53	54	55	56	57	58	59	60	61	62	
1935	52	53	54	55	56	57	58	59	60	61	These cohorts need to take up insurance in 1992 if they did not before because they planned to retire at, say 62
1936	51	52	53	54	55	56	57	58	59	60	
1937	50	51	52	53	54	55	56	57	58	59	
1938	49	50	51	52	53	54	55	56	57	58	
1939	48	49	50	51	52	53	54	55	56	57	
1940	47	48	49	50	51	52	53	54	55	56	
1941	46	47	48	49	50	51	52	53	54	55	
1942	45	46	47	48	49	50	51	52	53	54	
1943	44	45	46	47	48	49	50	51	52	53	
1944	43	44	45	46	47	48	49	50	51	52	
1945	42	43	44	45	46	47	48	49	50	51	
1946	41	42	43	44	45	46	47	48	49	50	
1947	40	41	42	43	44	45	46	47	48	49	
1948	39	40	41	42	43	44	45	46	47	48	
1949	38	39	40	41	42	43	44	45	46	47	
1950	37	38	39	40	41	42	43	44	45	46	
1951	36	37	38	39	40	41	42	43	44	45	
1952	35	36	37	38	39	40	41	42	43	44	These cohorts take up insurance as they turn 40 in order to get early retirement
1953	34	35	36	37	38	39	40	41	42	43	
1954	33	34	35	36	37	38	39	40	41	42	
1955	32	33	34	35	36	37	38	39	40	41	
1956	31	32	33	34	35	36	37	38	39	40	No early retirement motive
1957	30	31	32	33	34	35	36	37	38	39	
1958	29	30	31	32	33	34	35	36	37	38	
1959	28	29	30	31	32	33	34	35	36	37	
1960	27	28	29	30	31	32	33	34	35	36	
1961	26	27	28	29	30	31	32	33	34	35	
1962	25	26	27	28	29	30	31	32	33	34	
1963	24	25	26	27	28	29	30	31	32	33	

NOTE.— Ages of different cohorts over the years, and the corresponding early retirement regulations.

**Table A.3:** *Impact of 1992 Mortgage Reform on Unemployment Insurance Participation among Birth Cohorts 1960-1962 for Years 1987-1994*

	OLS	FE	LDV	FE Logit	LDV Logit
Continuous	-0.663** (0.127)	-0.585** (0.126)	-0.274** (0.0707)	-1.550** (0.489)	-0.122 (0.0637)
Observations (C)	421,096	421,096	368,459	52,688	368,336
Individuals (C)	52,637	52,637	52,637	6,586	52,618
Discrete	-0.751** (0.183)	-0.650** (0.181)	-0.362** (0.100)	-1.608* (0.825)	-0.212* (0.103)
Observations (D)	396,840	396,840	347,235	49,952	347,118
Individuals (D)	49,605	49,605	49,605	6,244	49,587

NOTE.—Standard errors clustered by individual in parentheses; \*  $p < 0.05$ , \*\*  $p < 0.01$ . The table collects Average Partial Effects (APE) on insurance up-take computed for the post-1991 subsample, in percentage points, from models of 0-1 unemployment insurance status estimated on the entire 1987-1995 panel. The continuous and discrete specifications correspond to Tables 1.2 and 1.3, respectively, restricted to those who had no speculative incentive to join an unemployment insurance fund in order to start collecting early retirement eligibility for age 60. See Section 1.3 for details of our sample selection otherwise. The model in the third column adds the lagged insurance status to control for inertia, and the second column shows coefficients from a fixed effect model. The fourth column shows results from a fixed effect logit model, estimated on switchers only. Finally, the model in the fifth column is a discrete choice model with fixed costs of switching, which translates into a lagged dependent variable logit. All models include the same controls used for the estimates shown in Tables 1.2 and 1.3.

Table A.4: Supplemental Security Income

	1987	1988	1989	1990	1991	1992	1993	1994 <sup>4</sup>	1995	1996
	After tax					Before tax				
<b>Basic amount (per month)</b>										
Single <sup>1</sup>	2579	2649	2728	2728	2796	2852	2909			
Couple	5158	5298	5456	5456	5592	5704	5818			
Breadwinner <sup>2</sup>								8852	8862	9057
Not breadwinner <sup>2</sup>								5546	6652	6803
Below age 23, non-breadwinner, living at home	1325	1361	1833.3	1857.9	1881	1826	1890	2080	2088	
Below age 23, non-breadwinner, not living at home	1847	1897	3000	3040.2	3135	3198	3310	4251	4268	
Below age 25, non-breadwinner, living at home										2138
Below age 25, non-breadwinner, not living at home										4370
<b>Child supplement for kids below age 18 (per year)<sup>3</sup></b>										
Proportion of normal benefits, oldest/first child	2	2	1.67	1.67	1.67	1.67	1.67			
Proportion of normal benefits, other kids	1.67	1.67	1.67	1.67	1.67	1.67	1.67			
Normal contribution per child	7176	7176	7392	7572	7764	7920	8076			
Amount, first child	14352	14352	12320	12620	12940	13200	13460			
Amount, other children	11960	11960	12320	12620	12940	13200	13460			
<b>Housing supplement</b>	Given on individual basis to cover housing expenses not already subsidized. Default was to subsidize all these expenses									

NOTE.— \* Benefit levels are per January 1, every year. Until 1993 benefit levels changed per July each year. After this, they changed January 1. \* The basic amounts of benefits are generally lower for young people under the age of 23 (25) until (after) 1995. Yet they received the full "adult" benefit level if they met certain working requirements: Until 1995 individuals under age 23 (21) would receive full benefits if they had worked consecutively for the 3 (12) months preceding unemployment at a standard wage rate (as opposed to the usually lower wage rate of young workers). After 1995 this rule applied to individuals under 25 who worked consecutively for the 12 months preceding unemployment (by April 1, 1995, the requirement became 18 months).

1 After a consecutive period of 9 months the amount decreased by DKK 300-400 (approx. DKK 300 in 1987 and approx. DKK 400 in 1996)

2 Married couples received the sum of their benefits.

3 If the applicant paid child support then the supplement would equal the amount of child support but could not exceed standard child-support contribution. Any child support received would be subtracted in the supplement.

4 From 1994 supplemental security income became taxable and the rules/rates changed (were simplified). Standard rate became 80% of unemployment insurance benefits for breadwinners and 50% for non-breadwinners (changes to 60% in 1995). Young people under age 23 with no breadwinner duties received less.

Sources: The series "Sociale Ydelser", Statistisk Årbog (DSt) samt "Arbejdsmarkedspolitisk Årbog"/"Arbejdsmarkedet og Arbejdsmarkedspolitik"

Table A.5: Unemployment Benefits and Basic Membership Fees

	1987	1988	1989	1990	1991	1992	1993	1994 <sup>4</sup>	1995	1996
<b>Maximum benefits (per day)<sup>1</sup></b>										
Full time insured	342	354	389	399	409	417	425	509	511	523
Part time insured <sup>2</sup>	228	236	259	266	273	278	283	339	341	349
Recent graduate, full-time insured <sup>3</sup>	271	283	311	319	327	334	340	417	419	429
Recent graduate, part-time insured	181	189	207	213	218	222	226	277	280	286
<b>Conversion</b>										
Day to week	6	6	6	6	6	6	6	5	5	5
Day to year	312	312	312	312	312	312	312	261	261	261
<b>Cost (per year, full-time insurance)</b>	2736	2832	3112	3192	3272	3336	3456	3552	3600	
<b>Bottom bracket MTR</b>	51.0%	51.6%	51.6%	51.5%	51.8%	52.1%	52.2%	47.5%	47.0%	47.1%

NOTE.— \* Benefit levels are per January 1, every year. Until 1993, benefit levels changed per July each year. After this, they changed January 1.

1 Benefit levels cannot exceed 90% of the average daily income calculated over the period of 12 weeks preceding the unemployment.

2 Maximum benefits for part time insured amounts to 2/3 of the full time insured benefit level.

3 The special benefit level for new graduates corresponds to 80% of the maximum standard benefits. (From 1994, 82%).

4 In 1994, the benefit period changed — effectively from infinity to 7 years.

Sources: The series "Sociale Ydelser", Statistisk Årbog (DSt) samt "Arbejdsmarkedspolitisk Årbog" / "Arbejdsmarkedet og Arbejdsmarkedspolitik"

## Appendix B

# Appendix to Chapter 2

### B.1 Supplementary Tables

**Table B.1:** *Summary Statistics for Experimental Sample*

Variable	Mean	Std. Dev.	Median
<i>Invitation to Referee (N = 2,423)</i>			
Agreed to submit review	66.2%	47.3%	
<i>Refereeing statistics conditional on agreement (N = 1,605)</i>			
Reviews censored (not submitted)	6.3%	24.3%	
Review time conditional on submitting review (days)	44.9	28.6	41.0
New referee (no historical data)	32.7%	46.9%	
<i>Referee Characteristics (N = 1,157)</i>			
Agreed to do 1 job during experiment	74.9%	43.4%	
Agreed to do 2 jobs during experiment	16.4%	37.1%	
Agreed to do 3+ jobs during experiment	8.6%	28.1%	
Tenured	54.6%	49.8%	
Academic	92.4%	26.5%	
American	52.5%	50.0%	
Female	12.3%	32.9%	

NOTE.—This table reports summary statistics for referee invitations sent between February 15, 2010 and May 9, 2011, the time period when the cash reward was offered. The first section of the table shows the fraction of referee requests that were accepted. The second section reports statistics for the subsample of referee requests that were accepted. A review is defined as censored if it is not submitted before the editor makes a decision on the paper. The summary statistics for review times are based on the subsample of submitted reviews. The third section of the table reports statistics on the referees who accepted the invitation and for whom the relevant information is available. See Appendix B.3 for the definitions of the variables used in this table.

**Table B.2: Randomization and Selection Tests**

Group:	6 Week	Social	4 Week	Cash	Equality test p-value
<i>A. Randomization Tests:</i>					
<i>Full Sample of All Invited Referees</i>					
Has pre-experiment data	58.2%	63.6%	66.0%	66.6%	0.07
Prior agreement rate	73.8%	70.3%	77.4%	73.8%	0.17
Prior median turnaround time	54.1	57.1	55.2	58.6	0.24
Tenured	60.2%	68.4%	59.8%	65.9%	0.07
Academic	90.2%	93.4%	93.0%	93.4%	0.51
American	53.4%	58.6%	53.8%	51.2%	0.30
Female	12.2%	8.3%	13.4%	11.8%	0.20
Observations	639	568	626	590	
<i>B. Selection Tests:</i>					
<i>Sample of Referees who Accepted Invitations</i>					
Has pre-experiment data	64.1%	65.1%	71.6%	68.2%	0.25
Prior agreement rate	82.5%	79.2%	87.3%	81.5%	0.03
Prior median turnaround time	52.1	57.1	53.8	57.0	0.19
Tenured	50.8%	59.9%	50.9%	59.4%	0.09
Academic	91.0%	96.2%	91.8%	93.0%	0.09
American	56.5%	57.9%	55.9%	51.1%	0.51
Female	14.1%	9.9%	16.1%	12.6%	0.30
Observations	432	347	401	425	

NOTE.—This table reports summary statistics of pre-experiment variables by treatment group. Panel A uses all referees invited to review a paper between February 15, 2010 and May 9, 2011 (the period when the cash reward was offered). Panel B replicates Panel A for the selected sample of referees who accepted the invitation to review. Column (5) reports the p-value for a test of equality of the coefficients across all four groups, clustering standard errors by referee (except for median review times). Has pre-experiment data is an indicator for having information in the editorial system at some point between November 1, 2005 and February 15, 2010, when the experiment began. Prior agreement rate is the fraction of reviews that the referee accepted during that period. Prior median review time is the median review time for the three most recent manuscripts reviewed before the experiment (among referees who reviewed manuscripts before the experiment). Tenured is an indicator for having tenure (based on CV's posted online) when the referee received the invitation; academic is an indicator for being in an academic position. American is an indicator for a US-based employer, and Female is a gender indicator from data collected manually. The number of observations in Panel A is the number of referee report invitations; in Panel B, it is the number of accepted invitations.

**Table B.3: Median Review Times by Treatment Group**

Sample	6 Week	Social	4 Week	Cash
Full Sample	47.8 (1.02) 432	45.9 (0.84) 347	35.5 (1.60) 401	27.5 (0.24) 425
Tenured Referees	50.4 (1.58) 203	46.8 (1.62) 199	44.1 (2.67) 189	27.7 (0.49) 236
Untenured Referees	45.9 (0.83) 197	45.5 (0.75) 133	31.7 (1.70) 182	27.3 (0.27) 161

NOTE.—This table shows the effects of the treatments on median review times. These estimates are reported in Figure 2.2a and Figure 2.4 and are reproduced here with standard errors as a reference. The sample includes all referees who accepted invitations sent between Feb. 15, 2010 and May 9, 2011 (the period when the cash reward was offered). Standard errors and number of observations are reported below each estimate. The first row of estimates uses the full-sample; the second and third rows restrict the sample to referees who were tenured vs. untenured at the time of the experiment. Tenure status was collected from CV's posted online and hence is not available for all referees. See Appendix B.3 for further details.

**Table B.4:** Cox Hazard Model Estimates of Treatment Effects on Review Times

	Extended Sample No Controls	Extended Sample with Controls	Primary Sample	Time-Varying Covariates
4 week deadline	0.266*** (0.0720)	0.393*** (0.0720)	0.418*** (0.0783)	0.391*** (0.0738)
Cash	0.388*** (0.0969)	0.502*** (0.0953)	0.485*** (0.0968)	0.161 (0.113)
Social	0.0769 (0.0637)	0.179** (0.0657)	0.152* (0.0746)	0.187** (0.0693)
Post-cash	0.185 (0.111)	0.255* (0.114)		0.242 (0.139)
Near deadline				0.864*** (0.0921)
Past deadline				1.188*** (0.0972)
Cash near deadline				1.007*** (0.169)
Cash past deadline				0.362 (0.216)
Post-cash near deadline				0.00595 (0.231)
Post-cash past deadline				-0.109 (0.269)
Controls		Yes	Yes	Yes
Number of spells	2,212	2,212	1,605	2,212

NOTE.—This table reports coefficients from Cox proportional hazard models, with standard errors clustered by referee in parentheses. The asterisks represent statistical significance: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ . The point estimates can be interpreted as the percentage impact of the variable on the baseline hazard rate (which measures hazards in the 6 week control group). Columns 1, 2 and 4 report estimates from the extended sample, which includes all invitations from February 15, 2010 to October 26, 2011. Column 3 reports estimates from the baseline sample, which includes invitations from February 15, 2010 to May 9, 2011 (the period during which the cash reward was offered). In all four columns, baseline hazards are stratified by invitation month and spells that last for more than 20 weeks (140 days) are censored at 140 days. The 4 week deadline indicator is 1 for both the four-week and cash incentive groups, who face four-week deadlines. The cash variable is an indicator for being in the cash incentive group while cash rewards were offered (prior to May 9); it is defined as 0 for all review invitations after May 9. The post-cash variable is an indicator for previously being in the cash incentive group; it is defined as 0 for all review invitations before May 9. The social variable is an indicator for being in the social incentive group. Columns 2-4 control for a referee's pre-experiment review times by including bimonthly indicator variables (up to 6 months) for the review time for each of the previous three referee reports. They also include controls for tenure, U.S. residence, working in academia, and the fraction of reviews the referee accepted prior to the start of the experiment at the *Journal of Public Economics*, as well as the number of words, tables, and equations in the article reviewed by the referee. See Appendix B.3 for definitions of all of these variables. Covariates are set to 0 if they are missing and all specifications include indicators for the observation having a missing value of the covariate. Column 4 includes terms allowing for time-varying hazard rates. Near and past deadline represent the period one week before and after the due date, respectively. These indicators are also interacted with the cash and post-cash indicators.

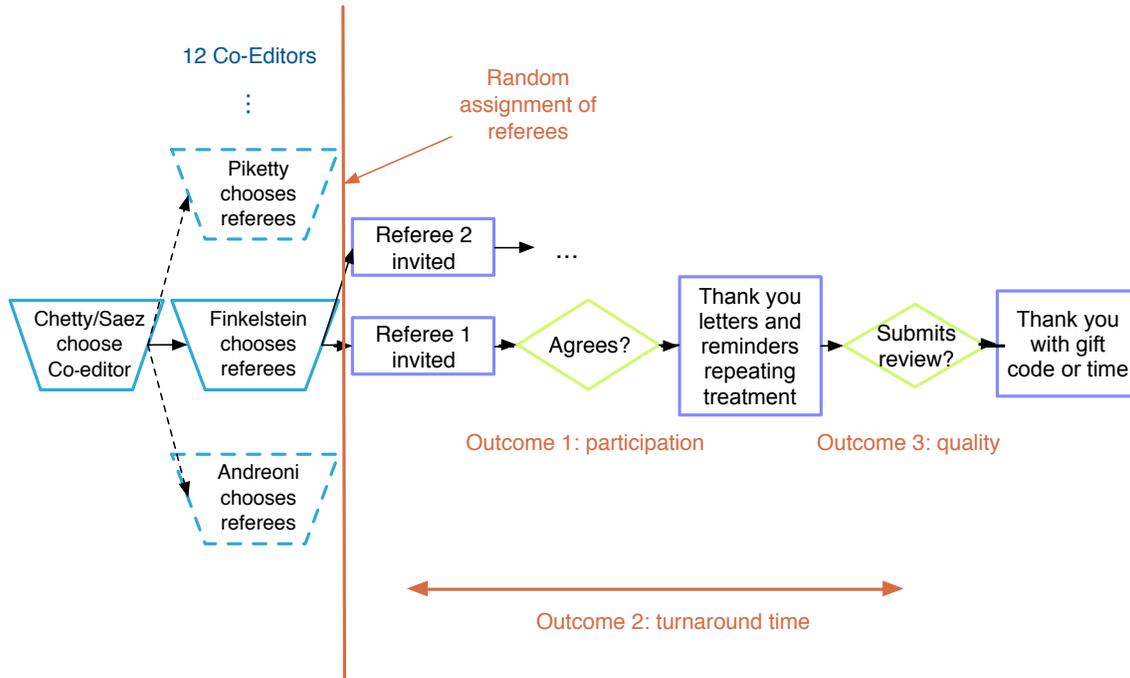
**Table B.5: Spillover Effects on Other Journals**

	6 Week	Social	4 Week	Cash
A. Reviewer Invitation Acceptance Rate at Other Journals				
Percent accepting invitation	62.1%	58.8%	60.6%	61.8%
	(2.31)	(2.55)	(2.38)	(2.26)
p-value for equality with control		0.344	0.654	0.702
Observations	999	806	969	993
B. Review Times at Other Journals				
Median review time (days)	56.2	54.0	56.5	57.0
	(1.39)	(1.72)	(1.81)	(1.71)
p-value for equality with control		0.562	0.596	0.894
Observations	620	474	587	614

NOTE.—This table reports estimates of spillover effects of the treatments on referee behavior at other Elsevier journals during the experimental period. The sample includes all referees who accepted invitations to review papers for the *Journal of Public Economics* between February 15, 2010 and May 9, 2011 (the period when the cash reward was offered). We use data from other Elsevier journals in related fields (see Appendix B.6 for a list) in this table, restricting attention to reviewer invitations received after the first invitation during the experimental period at the *Journal of Public Economics* and before December 31, 2011. In Panel A, the outcome is the percentage of referees who accept invitations to review papers at other journals. We report standard errors, clustered by referee, in parentheses. We also report p-values for the null hypothesis that the percentages are the same in each treatment group and its corresponding control group. For the social and 4 week groups, the control group is defined as the 6 week deadline group. For the cash incentive group, the control group is defined as the 4 week deadline group, which is the relevant comparison because the cash incentive group also faced a 4 week deadline. There is one observation for each review invitation that referees received from other Elsevier journals. In Panel B, the outcome is the median number of days taken to submit a review conditional on accepting the invitation to referee. Standard errors are reported in parentheses and the p-values are for hypothesis tests analogous to those in Panel A. The number of observations is the number of referee reports submitted to the other journals.

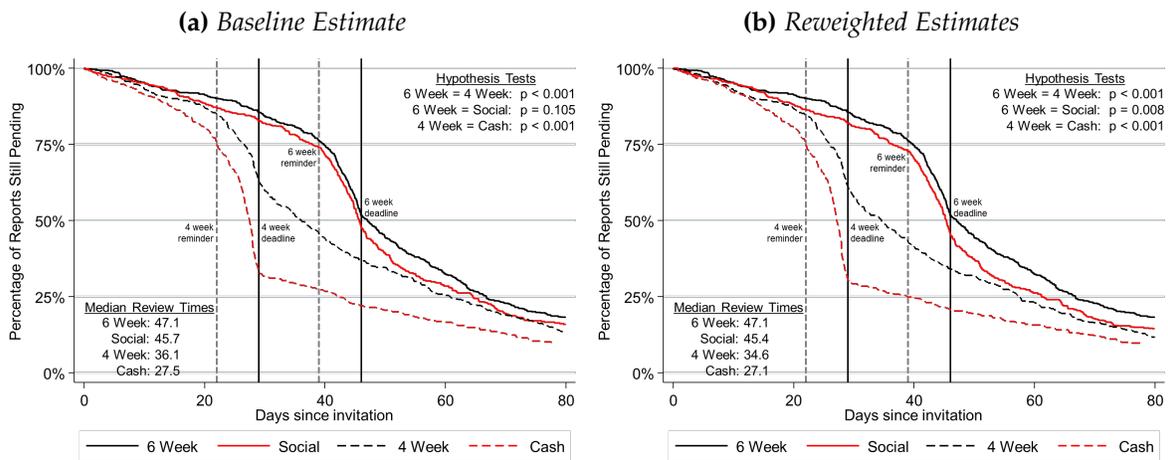
## B.2 Supplementary Figures

Figure B.1: *Timeline of Interventions and Outcomes*



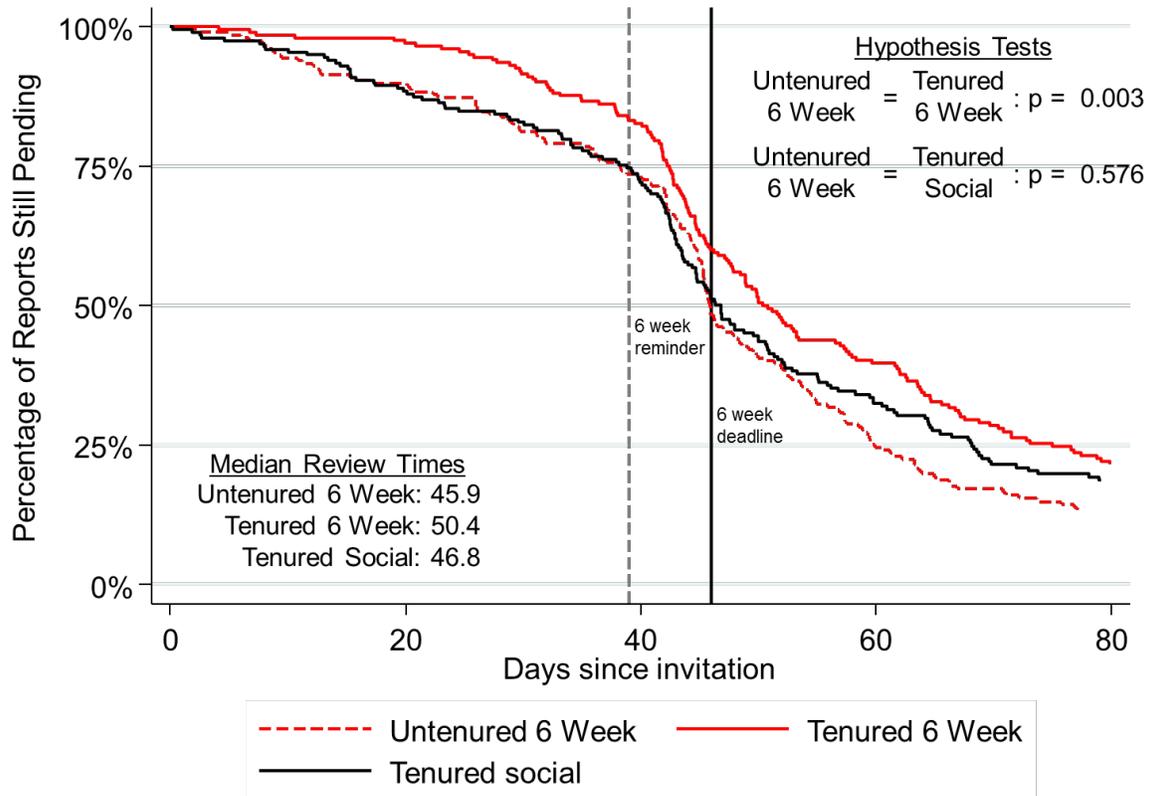
NOTE.—The figure depicts the timeline of the refereeing process during the experiment. Once a submission is received, editors assign a co-editor in charge who then chooses referees. Invited referees are randomly assigned to one of the four groups (six-week, four-week, cash, social) and receive an email invitation tailored to their group (shown in Appendix B.7). Referees accept or decline the invitation, which is the first outcome we study. If they accept, we send group-specific reminders one week before the deadline (shown in Appendix B.8). We then measure the time taken to submit a review, the second outcome we study. If a review is submitted, we send a thank you letter with the cash reward (to eligible referees) and measure the quality of the report, the third outcome we study.

**Figure B.2: Review Times by Treatment Group in Extended Sample**



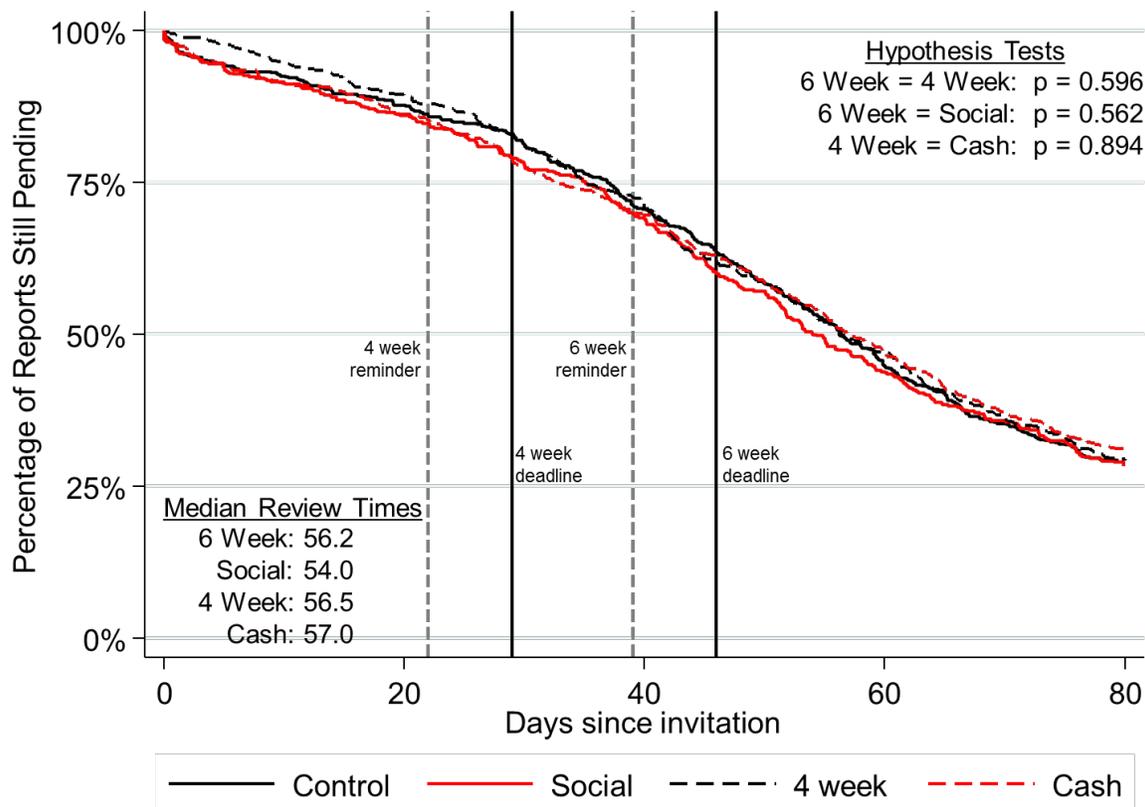
NOTE.—This figure replicates Figure 2.2 using the full experimental period from February 15, 2010 to October 26, 2011, including the period after May 9, when the cash reward was stopped. The cash group in this figure still includes only referee invitations up to May 9, 2011. The other groups include all invitations during the full experiment. See notes to Figure 2.2 for details on the construction of this figure and Appendix Table B.4 for Cox hazard model estimates corresponding to these survival curves.

**Figure B.3: Social Incentives and Tenured vs. Untenured Referees**



NOTE.—This figure plots a subset of the survival curves reported in Figure 2.4 on a single figure to show that tenured referees have longer turnaround times than untenured referees in the control group, but behave like untenured referees when facing social pressure. We replicate the series in Figure 2.4 for (a) untenured referees in the six-week group, (b) tenured referees in the six-week group, and (c) tenured referees in the social group. The solid vertical line depicts the six week deadline relevant for these groups. The dashed vertical line depicts the deadline reminders sent one week before this deadline. We report median review times, defined as the point at which the fraction of reports pending is 50 percent, for each group. We also report p-values from non-parametric Wilcoxon tests for the hypothesis that review times are the same in the untenured six-week group and the two tenured groups. We truncate the x-axis at 80 days in the figure for scaling purposes, but use all available data for the hypothesis tests.

**Figure B.4:** Spillover Effects: Review Times at Other Journals



NOTE.—This figure shows the effects of our experimental interventions at the *Journal of Public Economics* on referees’ review times at other Elsevier journals (listed in Appendix B.6). The sample includes all referees who accept a refereeing invitation at another Elsevier journal (before December 31, 2011) after receiving an invitation to referee at the *Journal of Public Economics* during our primary experimental period, February 15, 2010 to May 9, 2011. Each survival curve plots the percentage of reports still pending vs. the number of days elapsed since the referee received the invitation from the other journal. As a reference, the solid vertical lines depict the six week deadline (45 days) and the four week deadline (28 days) used at the *Journal of Public Economics* during the experiment. The dashed vertical lines depict the reminders sent one week before each deadline. Other journals have different deadlines and reminder policies. We report median review times, defined as the point at which the fraction of reports pending is 50 percent, for each group. We also report p-values from non-parametric Wilcoxon tests for the hypothesis that review times at other journals are the same in each treatment group and its corresponding control group. We compare the four-week and social incentive groups to the six week group. We compare the cash group to the four-week group because the cash group also faced a four week deadline. We truncate the x-axis at 80 days in the figure for scaling purposes, but use all available data for the hypothesis tests.

### **B.3 Data Sources and Variable Definitions**

*Data from the Journal of Public Economics.* Our primary source of data is the Elsevier online editorial system. We downloaded data from this system on July 22, 2012 for the analysis reported in the paper. We use data on all referees invited to review a new submission between February 15, 2010 and October 26, 2011. We exclude 15 observations that were contaminated (e.g. by letters with errors) and 5 observations in which the referee did not receive the email invitation. Referee requests for revisions are excluded from the experiment and are always subject to the 6 week deadline.

The Elsevier data system records time of invitation, agreement and submission of the report. Using these data, we generate an indicator for accepting the invitation, the turnaround time in days, and month of invitation. We obtain data on referees' review invitations and turnaround times prior to the start of the experiment from the same database, which contains information going back to November 1, 2005. The online system uses a single numerical identifier for a referee; we consolidate a few cases where referees have multiple ID's manually using the reviewer's name and affiliation.

The editorial system also stores all referee reports, which are submitted either as file attachments or as plain text via an online form. We define word counts for the referee reports as the sum of the words in the online text forms and any attachments. We use a similar procedure to measure the word count of manuscripts as well as the number of tables and equations in each manuscript. Note that these automated counting procedures do not always deliver accurate counts, but we expect such measurement error to be balanced across the treatment groups.

Each referee must select a recommendation for the manuscript on an online menu (accept, revise-and-resubmit, or reject). We use this information to define an indicator for whether the editor follows the referee's recommendation on whether or not to reject the submission, grouping the accept and revise-and-resubmit categories into a single category.

*Demographics.* We collected demographic information by locating referees' CVs online. We downloaded these CVs during Fall 2010, with an update for new referees in November

2011. We use these CVs to define indicators for gender, tenure status, working in the U.S., and working in an academic position. Tenure status is defined as being a full professor at a university or mentioning tenure on the CV for any other position. Working in the U.S. is based on the employer’s address and an academic position is defined as having an affiliation with a university. We code these variables as missing for referees for whom we were unable to locate CVs online or whose CVs did not contain the relevant information. We located CVs for 92.9% of the 1,606 referee reports in our primary (February 15, 2010 to May 9, 2011) sample.

*Data from Other Journals.* We obtain data from other Elsevier journals (listed in Appendix B.6) from the Elsevier editorial system. We compiled the longest histories available in the system for each journal. The available data vary across the journals, with the earliest records going back to November 2005. We use data up to December 31, 2011 from other journals. Elsevier does not use a unique identifier for referees across journals. We therefore linked referees to their performance at other journals based on email addresses (after extensive manual cleaning to match text fields).

## B.4 Reweighting Methodology

This appendix describes the reweighting procedure used to construct Figure 2.2b. We first discretize each referee’s most recent pre-experiment review time into eight bins,  $b = 1, \dots, 8$ : seven monthly indicators for the pre-experiment review time if available (<30, 30-59, 60-89, 90-119, 120-149, 150-179, and  $\geq 180$  days) and an indicator for having no pre-experiment data.

To reweight the social incentive group to match the six-week control group, we take the referees assigned to those two groups and calculate the fraction of observations in bin  $b$  in the social incentive group, which we denote by  $p_b$ . The fraction of observations in bin  $b$  in the six-week control group is  $1 - p_b$ . We weight each observation  $i$  by  $\frac{1 - p_{b(i)}}{p_{b(i)}}$  when estimating the survival curve for the social group, where  $b(i)$  denotes the bin to which observation  $i$  belongs.

We reweight the cash and four-week groups to match the control group on pre-experiment durations using the same approach. The survival curve for the six-week (control) group is unchanged by definition.

To adjust for differences in pre-experiment durations when testing for the equality of the survival curves, we conduct unweighted Wilcoxon tests that are stratified by the bin variable  $b$ .

## B.5 Hazard Model Estimates of Treatment Effects on Review Times

This appendix presents estimates of the impacts of the treatments on review times using Cox hazard models. Let  $h_{it}$  denote the hazard rate of submitting a referee report  $t$  days after the invitation (i.e., the probability of submitting a report on day  $t$  conditional on not submitting prior to day  $t$ ). Let  $\alpha_{mt}$  denote the baseline hazard rate for referees in the six-week control group who receive an invitation to review a paper in month  $m$  of the experiment. We stratify the baseline hazards by invitation month to account for any differences over time in referee behavior. The Cox hazard model specification is

$$h_{it} = \alpha_{mt} \exp(\beta_1 \text{fourweek}_i + \beta_2 \text{cash}_i + \beta_3 \text{social}_i + \beta_4 \text{postcash}_i + \gamma X_i)$$

In this specification, the *fourweek* indicator is 1 for both the four-week and cash incentive groups, who face four-week deadlines. Hence, the coefficients on the cash variables represent the effect of the cash treatment over and above the four-week deadline effect. The *cash* variable is an indicator for being in the cash incentive group while cash rewards were offered (prior to May 9); it is defined as 0 for all review invitations after May 9. The *social* variable is an indicator for being in the social incentive group. The *postcash* variable is an indicator for previously being in the cash incentive group; it is defined as 0 for all review invitations before May 9. The vector  $X_i$  is a set of controls that we vary across specifications. We censor spells that last for more than 20 weeks at 140 days to reduce the influence of outliers and we cluster standard errors by referee.

We report estimates from variants of this model in Appendix Table B.4. We begin in

Column 1 by estimating the hazard model with no additional controls (no  $X$  vector). We use the extended sample, which includes all invitations from February 15, 2010 to October 26, 2011, in this specification. Consistent with the results in Figure 2.2, we find that both the four week deadline and the cash incentive substantially increase hazard rates of submitting reports, i.e. reduce review times. The estimates  $\beta_j$  can be interpreted as the percentage impact of the variable on the baseline hazard rate. For example, the coefficient of 0.266 on the four-week indicator implies that the hazard rate is 26.6% higher on average for referees facing a four-week deadline relative to those facing a six-week deadline. The point estimate on the post-cash indicator is positive and marginally significant, supporting the view that there is no crowd-out of intrinsic motivation for referees who previously received cash incentives. The estimated impact of social incentives is small and statistically insignificant. This is consistent with Figure 2.2a, which shows that we do not detect significant differences between the social incentive and control groups when comparing raw distributions of review times.

Column 2 adds a rich set of controls for referee and manuscript characteristics to the specification in Column 1. We control for a referee's pre-experiment review times by including bi-monthly indicator variables (up to 6 months) for the review time for each of the previous three referee reports. We also include controls for tenure, working in the U.S., working in academia, and referees' agreement rates to invitations in the available history of the *Journal of Public Economics*, as well as the number of words, tables, and equations in the article reviewed by the referee. The covariates are set to 0 if they are missing and all specifications include indicators for the observation having a missing value of the covariate. Hence, the sample is exactly the same as in Column 1.

The inclusion of the controls increases the estimated impact of the social incentive treatment significantly. This result confirms the pattern in Figure 2.2b, showing that referees who agree to review manuscripts under the social incentive treatment are slightly negatively selected in terms of review times. Adjusting for these differences in pre-experiment turnaround times and other observables, we find that the social incentive treatment in-

creases hazard rates by approximately 18% relative to the six-week deadline. The cash and four-week treatments continue to have highly significant impacts on hazard rates with controls.

Column 3 replicates Column 2 restricting the sample to the primary experimental period from February 15, 2010 to May 9, 2011, when the cash reward was offered, as in the main text. We find that the impacts of the shorter deadline, cash incentives, and social incentives are all very similar when we restrict to this subset of referee reports.

The preceding specifications all assume that the treatments have a constant percentage impact on hazard rates throughout the spell. However, the non-parametric survival curves in Figure 2.2 show that this proportional hazards assumption is not a good approximation. In particular, the four week deadline and cash incentives have much greater effects before the deadline than after the deadline, as one would expect. To account for these responses, in Column 4 we estimate a Cox model that extends Column 2 to permit time-varying covariates. We include indicators for being near the deadline and past the deadline, which represent the period one week before and after the due date, respectively. We also interact these indicators with the cash and post-cash indicators to capture the greater impacts of the cash treatment before the deadline.

Consistent with the patterns in Figure 2.2, the time varying covariates are highly significant: hazard rates are 86% higher in the week before the deadline and 119% higher in the week after the deadline. The cash treatment increases hazard rates by 100% in the week before the deadline but does not have a statistically significant effect in the week after the deadline. The post-cash treatment has no time-varying effect, as one would expect. The estimated impact of the social incentive remains similar to the other specifications. Overall, the model with time-varying covariates confirms the results in Figure 2.2 and shows that all three treatments have significant effects on referee behavior.

## **B.6 List of Other Journals Used to Assess Spillover Effects**

- Economics & Human Biology

- Economics Letters
- Energy Economics
- European Economic Review
- European Journal of Political Economy
- Games and Economic Behavior
- Journal of Banking & Finance
- Journal of Comparative Economics
- Journal of Corporate Finance
- Journal of Development Economics
- Journal of Economic Behavior & Organization
- Journal of Economic Psychology
- Journal of Environmental Economics and Management
- Journal of Health Economics
- Journal of International Economics
- Journal of Monetary Economics
- Journal of Urban Economics
- Labour Economics
- Regional Science and Urban Economics
- Resource and Energy Economics

## **B.7 Invitation Emails**

**Figure B.5: Control and Four-Week Invitation Emails**

*Cash Invitation Email*

Subject: Reviewer Invitation from JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS

Dear REFEREE,

You are invited to review the above-mentioned manuscript for publication in the Journal of Public Economics. The manuscript's abstract is at the end of this email.

If you accept this invitation, I would be very grateful if you would return your review by **July 4, 2010** (4 weeks from now). As a token of appreciation for timely reviews, you will receive a **\$100 Amazon.com® Gift Card\*** if you submit your report before the due date. The Journal of Public Economics will automatically email you a gift card code within a day after we get your report (no paperwork required).

Please choose one of the following options to proceed:

- 1) If you are willing to review this manuscript, please click: [Agree to Review](#)
- 2) If you are not able to review this manuscript, please click: [Decline to Review](#)
- 3) If you would like to view the manuscript before making a decision, please click: [View Manuscript](#)

To assist you in the reviewing process, I am delighted to offer you full access to Scopus (the largest abstract and citation database of research information) for 30 days. With Scopus you can search for related articles, references and papers by the same author. You may also use Scopus for your own purposes at any time during the 30-day period. If you already use Scopus at your institute, having this 30 day full access means that you will also be able to access Scopus from home. Access instructions will follow once you have accepted this invitation to review.

Yours sincerely,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

*Social Invitation Email*

Subject: Reviewer Invitation from JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS

Dear REFEREE,

You are invited to review the above-mentioned manuscript for publication in the Journal of Public Economics. The manuscript's abstract is at the end of this email.

If you accept this invitation, I would be very grateful if you would return your review by **July 21, 2010** (6 weeks from now). In the interest of improving transparency and efficiency in the review process, Elsevier will **publish referee times by referee name**, as currently done by the Journal of Financial Economics at [this website](#). The referee times for reports received between Jan 1, 2010 and Dec 31, 2010 will be posted on the Journal of Public Economics website in January 2011. Note that **referee anonymity will be preserved as authors only know the total time from submission to decision (and not individual referee's times)**.

Please choose one of the following options to proceed:

- 1) If you are willing to review this manuscript, please click: [Agree to Review](#)
- 2) If you are not able to review this manuscript, please click: [Decline to Review](#)
- 3) If you would like to view the manuscript before making a decision, please click: [View Manuscript](#)

To assist you in the reviewing process, I am delighted to offer you full access to Scopus (the largest abstract and citation database of research information) for 30 days. With Scopus you can search for related articles, references and papers by the same author. You may also use Scopus for your own purposes at any time during the 30-day period. If you already use Scopus at your institute, having this 30 day full access means that you will also be able to access Scopus from home. Access instructions will follow once you have accepted this invitation to review.

Yours sincerely,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

## Figure B.6: Cash and Social Invitation Emails

### *Control Group Reminder Email*

Subject: Reminder to review for JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS  
Journal of Public Economics

Dear REFEREE,

Thank you for agreeing to review this manuscript for the JPubE. I am writing to remind you that I would appreciate receiving your review by **July 21, 2010**, in a week.

You may submit your comments online in our editorial system by clicking [here](#). Please login as a Reviewer using the username and password I sent you in my first email.

You may access the manuscript by selecting the "Pending Assignments" link on your Main Menu page. To submit your comments, please click on the "Submit Reviewer Recommendation" link.

With kind regards,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

### *Four-Week Deadline Reminder Email*

Subject: Reminder to review for JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS  
Journal of Public Economics

Dear REFEREE,

Thank you for agreeing to review this manuscript for the JPubE. I am writing to remind you that I would appreciate receiving your review by **July 4, 2010**, in a week.

You may submit your comments online in our editorial system by clicking [here](#). Please login as a Reviewer using the username and password I sent you in my first email.

You may access the manuscript by selecting the "Pending Assignments" link on your Main Menu page. To submit your comments, please click on the "Submit Reviewer Recommendation" link.

With kind regards,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

Subject: Reminder to review for JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS  
Journal of Public Economics

## B.8 Reminder and Thank-You Emails

**Figure B.7:** *Control and Four-Week Reminder Emails*

### *Cash Incentive Reminder Email*

Dear REFEREE,

Thank you for agreeing to review this manuscript for the JPubE. I am writing to remind you that I would appreciate receiving your review by **July 4, 2010**, in a week. As a token of gratitude for timely reviews, you will receive a **\$100 Amazon.com® Gift Card\*** if you submit your report before the due date. The Journal of Public Economics will automatically email you a gift card code within a day after we get your report (no paperwork required).

You may submit your comments online in our editorial system by clicking [here](#). Please login as a Reviewer using the username and password I sent you in my first email.

You may access the manuscript by selecting the "Pending Assignments" link on your Main Menu page. To submit your comments, please click on the "Submit Reviewer Recommendation" link.

With kind regards,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

### *Social Incentive Reminder Email*

Subject: Reminder to review for JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS  
Journal of Public Economics

Dear REFEREE,

Thank you for agreeing to review this manuscript for the JPubE. I am writing to remind you that I would appreciate receiving your review by **July 21, 2010**, in a week. In the interest of improving transparency and efficiency in the review process, Elsevier will **publish referee times by referee name**, as currently done by the Journal of Financial Economics at [this website](#). The referee times for reports received between Jan 1, 2010 and Dec 31, 2010 will be posted on the Journal of Public Economics website in January 2011. Note that referee anonymity will be preserved as authors only know the total time from submission to decision (and not individual referee's times).

You may submit your comments online in our editorial system by clicking [here](#). Please login as a Reviewer using the username and password I sent you in my first email.

You may access the manuscript by selecting the "Pending Assignments" link on your Main Menu page. To submit your comments, please click on the "Submit Reviewer Recommendation" link.

With kind regards,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

## Figure B.8: Cash and Social Reminder Emails

### *Control Group and Four Week Deadline Thank You Email*

Subject: Reminder to review for JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS  
Journal of Public Economics

Dear REFEREE,

Thank you for your review of this manuscript.

You may access your review comments and the decision letter (when available) by logging onto the [Elsevier Editorial System](#). Please login as a Reviewer.

Kind regards,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

w

### *Cash Incentive Thank You Email*

Subject: Reminder to review for JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS  
Journal of Public Economics

Dear REFEREE,

Thank you for your review of this manuscript. As a token of appreciation for timely reviews, here is your \$100 Amazon.com® Gift Card\* code: [Claim Code](#). You are able to use it any time to make purchases at [Amazon.com](#) without any paperwork. If you experience any problems with it, please do not hesitate to contact me at [jpubec@gmail.com](mailto:jpubec@gmail.com).

You may access your review comments and the decision letter (when available) by logging onto the [Elsevier Editorial System](#). Please login as a Reviewer.

Kind regards,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

### *Social Incentive Thank You Email*

Subject: Reminder to review for JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS  
Journal of Public Economics

Dear REFEREE,

Thank you for your review of this manuscript.

You may access your review comments and the decision letter (when available) by logging onto the [Elsevier Editorial System](#). Please login as a Reviewer.

As you may remember, Elsevier will **publish referee times by referee name**, as currently done by the Journal of Financial Economics at [this website](#). Your time of 27 days for this review will be posted on the Journal of Public Economics website in January 2011. Note that referee anonymity will be preserved as authors only know the total time from submission to decision (and not individual referee's times).

Kind regards,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

## Figure B.9: Thank-You Emails

### Cash Invitation Email

Subject: Reviewer Invitation from JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS

Dear REFEREE,

You are invited to review the above-mentioned manuscript for publication in the Journal of Public Economics. The manuscript's abstract is at the end of this email.

If you accept this invitation, I would be very grateful if you would return your review by **July 4, 2010** (4 weeks from now). As a token of appreciation for timely reviews, you will receive a **\$100 Amazon.com® Gift Card\*** if you submit your report before the due date. The Journal of Public Economics will automatically email you a gift card code within a day after we get your report (no paperwork required).

Please choose one of the following options to proceed:

- 1) If you are willing to review this manuscript, please click: [Agree to Review](#)
- 2) If you are not able to review this manuscript, please click: [Decline to Review](#)
- 3) If you would like to view the manuscript before making a decision, please click: [View Manuscript](#)

To assist you in the reviewing process, I am delighted to offer you full access to Scopus (the largest abstract and citation database of research information) for 30 days. With Scopus you can search for related articles, references and papers by the same author. You may also use Scopus for your own purposes at any time during the 30-day period. If you already use Scopus at your institute, having this 30 day full access means that you will also be able to access Scopus from home. Access instructions will follow once you have accepted this invitation to review.

Yours sincerely,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

### Social Invitation Email

Subject: Reviewer Invitation from JPubE  
Ref. No.: JPUBE-D-10-00001  
Title: TITLE  
Editor: CO-EDITOR  
Author(s): AUTHORS

Dear REFEREE,

You are invited to review the above-mentioned manuscript for publication in the Journal of Public Economics. The manuscript's abstract is at the end of this email.

If you accept this invitation, I would be very grateful if you would return your review by **July 21, 2010** (6 weeks from now). In the interest of improving transparency and efficiency in the review process, Elsevier will **publish referee times by referee name**, as currently done by the Journal of Financial Economics at [this website](#). The referee times for reports received between Jan 1, 2010 and Dec 31, 2010 will be posted on the Journal of Public Economics website in January 2011. Note that **referee anonymity will be preserved as authors only know the total time from submission to decision (and not individual referee's times)**.

Please choose one of the following options to proceed:

- 1) If you are willing to review this manuscript, please click: [Agree to Review](#)
- 2) If you are not able to review this manuscript, please click: [Decline to Review](#)
- 3) If you would like to view the manuscript before making a decision, please click: [View Manuscript](#)

To assist you in the reviewing process, I am delighted to offer you full access to Scopus (the largest abstract and citation database of research information) for 30 days. With Scopus you can search for related articles, references and papers by the same author. You may also use Scopus for your own purposes at any time during the 30-day period. If you already use Scopus at your institute, having this 30 day full access means that you will also be able to access Scopus from home. Access instructions will follow once you have accepted this invitation to review.

Yours sincerely,

Liz Anderson  
Senior Editorial Assistant  
Journal of Public Economics

## B.9 Summary of Appendix Tables and Figures

Appendix Figure B.1 depicts the timeline of the refereeing process during the experiment.

Appendix Figure B.2 replicates Figure 2.2 from the text using the full experimental period from February 15, 2010 to October 26, 2011. This figure includes the period after May 9, when the cash reward was stopped, for the four week, six week, and social incentive groups. For the cash group, we continue to use data only up to May 9.

Appendix Figure B.3 plots a subset of the survival curves reported in Figure 2.4 in the main text on a single figure to show that tenured referees have longer turnaround times than untenured referees in the control group, but behave like untenured referees when facing social pressure.

Appendix Figure B.4 shows survival curves for review times at other Elsevier journals by the treatment group to which referees were assigned at the *Journal of Public Economics*.

Appendix Table B.1 presents the summary statistics for referee invitations sent between February 15, 2010 and May 9, 2011, the time period when the cash reward was offered.

Panel A of Appendix Table B.2 presents randomization tests for the set of referees invited during the primary experimental period. Panel B replicates Panel A in the subsample of referees who accept the invitations to test for selection effects.

Appendix Table B.3 presents estimates of treatment effects on median review times.

Appendix Table B.4 presents Cox hazard model estimates of the effects of the treatments on review times.

Appendix Table B.5 reports estimates of the effects of the treatments at the *Journal of Public Economics* on acceptance rates and review times at other Elsevier journals during the experimental period.

## Appendix C

# Appendix to Chapter 3

### C.1 Proof of Proposition

*Proof.* Denote by  $p_t$  the probability measure over  $\mathcal{S}(t)$  induced by  $p$ . We first show that  $p_t$  is pinned down by  $\succsim$ . Write  $B_t = \{b^0, \dots, b^N\}$  such that  $b^0 \leq \dots \leq b^N$  (the  $t$  subscripts are dropped to ease notation). For any  $i \leq N$ , let  $x_{m,s^t}$  be the reward that yields prize  $m$  at time  $t$  conditional on  $s^t = (s_1, \dots, s_{t-1}, b^i)$ . Denote by  $\psi(x_{m,s^t})$  the reward that yields a prize immediately such that  $\psi(x_{m,s^t}) \sim x_{m,s^t}$ .<sup>1</sup> Identify the immediate prize with  $\psi(x_{m,s^t})$ . The representation implies that

$$u(b^* + \psi(x_{m,s^t})) - u(b^*) = D(t)p_t(s_1, \dots, s_{t-1}, b^i)[u(b^i + m) - u(b^i)].$$

Note that since  $u$  is a strictly increasing diffeomorphism,  $\psi(x_{m,s^t})$  is a strictly increasing differentiable function of  $m$  that takes the value 0 when  $m = 0$ . Taking a derivative of the above expression with respect to  $m$  yields

$$u'(b^* + \psi(x_{m,s^t})) \frac{\partial \psi(x_{m,s^t})}{\partial m} = D(t)p_t(s_1, \dots, s_{t-1}, b^i)u'(b^i + m).$$

---

<sup>1</sup>The existence of such a reward is implied by the continuity of  $u$  together with either the assumption that  $D \leq 1$  or that  $u$  is unbounded. Its uniqueness is implied by the strict monotonicity of  $u$ .

Evaluating at  $m = 0$  gives

$$\frac{\partial \psi(x_{m,s^t})}{\partial m} \Big|_{m=0} = D(t) p_t(s_1, \dots, s_{t-1}, b^i) \frac{u'(b^i)}{u'(b^*)}. \quad (\text{C.1})$$

Consider the contingent reward  $x_{m,s^t}^{b^1-b^0}$  that gives  $b^1 - b^0$  at time  $t$  unconditionally and in addition gives  $m$  at time  $t$  conditional on  $s^t = (s_1, \dots, s_{t-1}, b^0)$ . Denote by  $\psi(x_{m,s^t}^{b^1-b^0})$  the contingent prize that pays  $b^1 - b^0$  at time  $t$  unconditionally and also an immediate prize (which, abusing notation, is also denoted by  $\psi(x_{m,s^t}^{b^1-b^0})$ ) satisfies  $\psi(x_{m,s^t}^{b^1-b^0}) \sim x_{m,s^t}^{b^1-b^0}$ . Then by the representation,

$$\begin{aligned} u(b^* + \psi(x_{m,s^t}^{b^1-b^0})) - u(b^*) &= D(t) p_t(s_1, \dots, s_{t-1}, b^0) [u(b^0 + b^1 - b^0 + m) - u(b^0 + b^1 - b^0)] \\ &= D(t) p_t(s_1, \dots, s_{t-1}, b^0) [u(b^1 + m) - u(b^1)]. \end{aligned}$$

Moreover,

$$\frac{\partial \psi(x_{m,s^t}^{b^1-b^0})}{\partial m} \Big|_{m=0} = D(t) p_t(s_1, \dots, s_{t-1}, b^0) \frac{u'(b^1)}{u'(b^*)}.$$

Similarly, for each  $0 < i \leq N$ ,

$$\frac{\partial \psi(x_{m,s^t}^{b^i-b^{i-1}})}{\partial m} \Big|_{m=0} = D(t) p_t(s_1, \dots, s_{t-1}, b^{i-1}) \frac{u'(b^i)}{u'(b^*)}. \quad (\text{C.2})$$

By assumption,  $p_t(s^t) \neq 0$  for all  $s^t$  (that is,  $\frac{\partial \psi(x_{m,s^t})}{\partial m} \Big|_{m=0} \neq 0$  for all  $s^t$ ). Using (C.1) and (C.2), we get

$$\frac{\frac{\partial \psi(x_{m,s^t}^{b^i-b^{i-1}})}{\partial m} \Big|_{m=0}}{\frac{\partial \psi(x_{m,s^t}^{b^1-b^0})}{\partial m} \Big|_{m=0}} = \frac{p_t(s_1, \dots, s_{t-1}, b^{i-1})}{p_t(s_1, \dots, s_{t-1}, b^i)}$$

for all  $i$ . That is, the noted relative probabilities are determined uniquely by preferences. Conclude that  $p_t(s^t)$  is uniquely determined for each  $s^t$ . In particular,  $p$  is uniquely determined.

To see that  $D$  is uniquely pinned down by preferences, note that by (C.1), for  $b^i = b^*$  (that is,  $s^t = (s_1, \dots, s_{t-1}, b^0)$ ),

$$\frac{\partial \psi(x_{m,s^t})}{\partial m} \Big|_{m=0} = D(t) p_t(s^t). \quad (\text{C.3})$$

Thus, the uniqueness of  $p_t(s^t)$  implies that of  $D(t)$ .

□

## C.2 Supplementary Tables

**Table C.1: Robustness to Selection on Observables and Unobservables**

	Stable income after 2010 <sup>†</sup>	Stable income before and after 2010 <sup>††</sup>	New job less likely than 20% <sup>†††</sup>	Income rises, then plateaus <sup>††††</sup>
Without controls	Difference in Time Preference Variables (with Robust SEs)			
$D_u(1, 2) - D_u(0, 1)^*$	-0.13 (0.05)	-0.11 (0.07)	-0.07 (0.05)	0.10 (0.10)
$\beta_u^{**}$	0.11 (0.05)	0.09 (0.07)	0.06 (0.05)	-0.11 (0.10)
$\delta_u^{***}$	0.00 (0.03)	0.00 (0.04)	-0.03 (0.03)	-0.03 (0.06)
With controls	Difference in Time Preference Variables (with Robust SEs)			
$D_u(1, 2) - D_u(0, 1)^*$	-0.12 (0.05)	-0.09 (0.07)	-0.08 (0.06)	0.12 (0.09)
$\beta_u^{**}$	0.10 (0.05)	0.07 (0.07)	0.07 (0.06)	-0.12 (0.10)
$\delta_u^{***}$	0.00 (0.03)	0.00 (0.04)	-0.04 (0.03)	-0.04 (0.06)
Bound and $R_{max}^2$	Oster (2014) Bounds (and Maximal $R^2$ s)			
$D_u(1, 2) - D_u(0, 1)^*$	0.03 (0.80)	0.24 (0.32)	-0.23 (0.44)	0.38 (0.36)
$\beta_u^{**}$	-0.09 (0.55)	-0.34 (0.22)	0.24 (0.32)	-0.25 (0.82)
$\delta_u^{***}$	0.04 (0.08)	0.14 (0.02)	-0.35 (0.08)	-0.27 (0.13)
N	31	10	32	11

NOTE.—The table reports contrasts of time-preference variables between subsamples and their complements. These are implemented as OLS regressions of the respective measure on the group identifier, with White-Huber robust standard errors in parentheses. Uncontrolled, univariate regressions in the top pane can be compared with how the coefficient of interest changes in models with controls for the demographic variables listed in Table 3.1 in the middle pane. The bottom pane reports the bound on the contrast from Oster (2014), with the minimal  $R^2$  with unobservables necessary to make the true contrast zero (but the naive contrast confounded by selection on unobservables). See the main text for more explanation (also for why Oster calls this the maximal  $R^2$ , not a minimum). The controls include the covariates with missing values replaced with zeros, but add a separate indicators for each covariate's originally missing observations. \*Reporting the average difference between each subject's revealed uncompensated discount factor over the second year and the one over the first year. \*\*Reporting the average ratio of discount factors over the second year and over the first year. \*\*\*Reporting the average ratio of the two-year discount factor and the one-year discount factor. <sup>†</sup>Only including those whose realized annual income remained within 10% of their 2010 income both in 2011 and 2012. <sup>††</sup>Only including those whose annual incomes in 2008 and 2009 as well. <sup>†††</sup>Including those whose self-report the probability of getting a new job in either year after the experiment to be less than 20%. <sup>††††</sup>Including those whose realized annual income rose more than 10% in 2011 but less than 10% on top of that in 2012.