8-2016

Three essays in economics

James R. Bland
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By JAMES R. BLAND

Entitled
THREE ESSAYS IN ECONOMICS

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

TIMOTHY N. CASON
Chair

JOHN M. BARRON

JUSTIN L. TOBIAS

SANDRA MAXIMIANO

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): TIMOTHY N. CASON

Approved by: JOHN M. BARRON 6/28/2016

Head of the Departmental Graduate Program Date

THREE ESSAYS IN ECONOMICS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

James R. Bland

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2016

Purdue University

West Lafayette, Indiana

For Ardian and Jamie.

ACKNOWLEDGMENTS

Producing this kind of work is never an individual effort. I would therefore like to thank the following people for thoughtful advice, assistance, and direction that led to the completion of this dissertation, my ongoing sanity, and the planned work that will hopefully follow.

To my thesis committee Tim Cason (chair), Jack Barron, Sandra Maximiano, and Justin Tobias, thank you for being always approachable, but also letting me go off on my own tangents unchecked for weeks at a time (there are sections in here that I can attribute to a strict subset of those). Thank you for not letting those weeks turn into months. Your comments along the way have helped me see the big picture of where I'm going and how my work fits with more than I initially think of.

Nikos Nikiforakis is largely responsible for my conversion from a career in Engineering to Economics, my pursuit of a Ph.D., and to do one in the US. He has also frequently volunteered advice ranging from minute details of a paper to how I should market myself as an economist. While I think I would have enjoyed being an engineer, I am glad to have made the switch.

I am grateful for The Purdue Experimental Lunch group's ongoing advice from the beginning to the end of the first two chapters, as well as one failed attempt at a field experiment.

I thank attendees of the 2015 Southern Economics Association meetings and the 2016 Midwest Economics Association meetings for comments related to Chapter 1, as well as numerous attendees of my job market presentations.

I thank the attendees of the ExperiMetrix workshop in Alicante, where I presented an early version of Chapter 2, for advice in the early stages of this project. Anna Conte, Peter Moffatt, and Guillaume Fréchette were especially helpful. Yaroslav

Rosokha, Bob Slonim, and Claudio Mezzetti have also provided advice and offered extensions relating to this paper.

Chapter 3 would not have been possible without Jack Cook's expert advice and insight about the US health insurance market. Steve Martin provided many useful comments on this chapter.

All three chapters of this dissertation may not have been finished without reliable and fast computing. This was made possible by the Krannert Computing Center. Special thanks go to Jeff Ello and Ryan Drown for setting Amanda and me up with the beast we used for Chapter 3. You have been very accommodating.

Thanks go to the junior faculty in the economics department for invaluable advice about the job market. You were not asked to volunteer this (I'm assuming), but you went over the top in helping us market ourselves and improve our presentations.

Over my four years at Purdue, I have been delighted to share offices with Sebastian Linde, Andy Meisner, Warren Zhang, Jacob Brindley, Peter Wagner, and Kory Garner. We struck a good balance of quiet work and discussion.

Thanks to my brewing buddy Paul Thomas. Our brewing sessions frequently led to better understanding of both the Ph.D. program and American culture.

Finally, I thank my family for their ongoing support and and encouragement. To Margaret, Jim, and Emma, thanks for keeping me same with by-weekly Skype calls for the first year in West Lafayette, for proof-reading my instructions, and all the other things you do that are much more important. To Amanda, thank you for tolerating my over-exuberance about minute details of this dissertation, helping me better understand myself, and showing me that running is not so bad. To Ardian and Jamie, thanks for the hugs.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

ABSTRACT

Bland, James R. PhD, Purdue University, August 2016. Three Essays in Economics. Major Professor: Timothy Cason.

How Many Games Are We Playing? An Experimental Analysis of Choice Bracketing in Games

A subject brackets two decisions if she "choose[s] an option in each case without full regard to the other" (Rabin and Weizsacker, 2009). Although in most situations such behavior is unlikely to be optimal, it is well documented in experiments where subjects make decisions in the absence of strategic considerations. This paper uses an economic experiment to investigate whether subjects also bracket their decisions in games. Subjects played two Volunteer's Dilemmas at the same time, with the payoffs from both games added to their earnings. In a lottery task, subjects were generally revealed to be risk-averse narrow bracketers. Aggregate play in the Roommate's Dilemma is not consistent with predictions made by assuming all subjects either narrowly or broadly bracket. On the individual level, structural modeling suggests that most subjects bracket narrowly in the game.

JEL codes: C91, C92, D03, D81

Mixture models of behavior and nuisance parameters: a semi-parametric Bayesian approach (with Justin Tobias)

When there is more than one model of decision-making that could explain behavior in experiments, the mixture model is a useful tool in taking theory to data. The estimation results can inform the researcher about the prevalence of each model in the sample, and whether observable characteristics of subjects are predictors of which model they use. Each model typically specifies a function describing behavior,

but also requires individual-level "nuisance parameters" that must also be estimated. We demonstrate that restrictive econometric assumptions made on these individual parameters can result in the researcher overstating the importance of type heterogeneity (subjects using different decision rules), when in fact the cause of heterogeneous choices is subject parameter heterogeneity (subjects having different nuisance parameters). We propose a less restrictive assumption, and demonstrate its implementation on some existing experimental data.

JEL codes: C11, C14, C51, D81

Keywords: Mixture models

HOSPITAL-INSURER BARGAINING POWER AND NEGOTIATED RATES (with Amanda Cook)

In addition to risk-sharing, U.S. health insurance companies negotiate rates for services with hosptials. The price of service can vary depending on which entity, if any, is insuring the patient. Insurers (and possibly their customers) benefit from negotiating through lower prices, while hospitals benefit through higher patient volume.

Using Massachusetts' Center for Health Information and Analysis (CHIA) data, we use hospital and insurer characteristics to estimate negotiated prices specific to hospital-insurer pairs. We investigate the relationship between two important quantities: (i) the charged amount that hospitals bill for their services, and (ii) the amount that hospitals are paid for insured patients. These numbers differ because the former is a function only of the services provided and the hospital's "chargemaster" prices, while the latter is the result of negotiation.

We find that payments for privately insured patients are on average 38% of charges when payments are made on a fee-for-service basis. However this ratio varies greatly by hospital and insurer. Compared to community hospitals without an emergency room, academic medical centers are compensated 15% more for their services, and hospitals with an emergency room are compensated 7% more than those without.

1. HOW MANY GAMES ARE WE PLAYING? AN EXPERIMENTAL ANALYSIS OF CHOICE BRACKETING IN GAMES

A decision-maker who *narrowly brackets* and "faces multiple decisions tends to choose an option in each case without full regard to the other decisions and circumstances that she faces" (Rabin and Weizsacker, 2009). In some circumstances it is obvious that there are benefits to making decisions jointly, and it relatively easy to make them jointly. For example, one might aim to choose their favorite *combination* of shirt and pants rather than their favorite shirt and their favorite pants; or one may aim to choose the best combination of food and drink on a restaurant's menu, thus avoiding the possibility of consuming fish with red wine.[1] Other times, either the benefits of jointly optimizing are not so obvious; or they are clear to the decision-maker, but the additional cognitive cost of optimizing jointly, or *broadly bracketing*, outweighs the additional benefit. Tversky and Kahneman (1981), for example, demonstrate that a significant fraction of individuals can be made to forgo certain amounts of money when faced with paired lottery choice tasks. Indeed, it is rare that single decisions can be made optimally without considering how they interact with the consequences of other decisions. We are also frequently in situations where we need to consider the actions taken by others, and hence we also need to think strategically about our actions. This paper will analyze the intersection of the two aspects of decision-making outlined above: (1) choice bracketing (Read et al., 1999), which involves ignoring one decision while making another, and (2) strategic considerations.

To fix ideas, consider the following dilemma faced by two roommates. On a particular night, there are two tasks that need to be done in the apartment: (1) Taking

[1] Or as a colleague recalled, ordering the unsurprisingly underwhelming pairing of fish and Coca-Cola while at a job market dinner.

Table 1.1.: The Roommate's Dilemma (two Volunteer's Dilemmas). Any risk-neutral mixed-strategy Nash equilibria of this game must satisfy $(p_D, p_T) = (0.1, 0.35)$.

	D	nD
D	10 10	100 10
nD	10 100	0 0

	T	nT
T	56 56	160 56
nT	56 160	0 0

out the trash (action T) is critical, because if it is not done tonight the trash bags will sit in the apartment for a week and create an unpleasant smell; and (2) washing the dishes (D) is less critical because they can always be washed tomorrow morning, but both roommates enjoy a tidy apartment, and so would prefer for the dishes to be done tonight. Of the two tasks, taking out the trash is the more unpleasant. Such a situation can be modeled by the "games"[2] presented in Table 1.1. Each player can either do the dishes (D) or not do the dishes (nD), and they can also take out the trash (T) or not (nT). If a player is not risk neutral, once her opponent's strategy is known, optimization requires not only that she optimally choose whether to do the dishes and whether to take out the trash, but instead *which pair of actions* she should take. The decision to take out the trash cannot be optimally made in isolation because for a subset of strategies that the opponent could play, some action pairs result in distributions of payoffs that are second-order stochastically dominated by other choices, and so would not be optimal for any risk-averse player. For example, if she believes that her opponent will do the dishes with probability 20%, and take out the trash with probability 40%, then not only is the action pair $\{D, nT\}$ second-order stochastically dominated by playing $\{nD, T\}$, but she would only realize this if she considered both decisions jointly; that is, if she broadly bracketed.[3] Each "game" in isolation is a Volunteer's Dilemma[4] which can be used to model the decision to veto an unpopular proposal (e.g. Holt, 2006). Therefore if one is looking for a more

[2]Here I use quotes around "games" because game theory treats this situation as *one game*. They get treated as separate "games" when players narrowly bracket.
[3]This result is derived in Section A.1.3.
[4]See Diekmann (1985) for theoretical analysis, and Diekmann (1986) for experimental analysis.

serious example, one could instead think of the Roommate's Dilemma as modeling the decision of *which combination* of unpopular proposals to veto.

Whether choice bracketing is relevant in strategic settings is an empirical question, and this paper uses an economic experiment to investigate this. The Roommate's Dilemma is used as a testbed for bracketing behavior in games. I vary payoffs between three treatments such that assumptions of narrow and broad bracketing make different comparative static predictions. In fact, neither narrow nor broad bracketing predict the observed treatment effects well at the aggregate level. I then elicit subjects' risk preferences and bracketing behavior in a lottery choice task, where strategic considerations are absent. Structural estimates suggest that the majority of subjects narrowly bracket both in the Roommate's Dilemma and the lottery task, however bracketing behavior in one task does not predict behavior in the other.

1.1 Related literature

A set of choices are bracketed together when the decision-maker takes into account the interaction of these choices, but ignores the effect of choices outside of this set (Read et al., 1999). An agent *broadly brackets* if she considers the interactions of all choices within her choice set, and narrowly brackets if she ignores some of these interactions.

Narrow bracketing is well documented in lottery choice experiments by Tversky and Kahneman (1981) and Rabin and Weizsacker (2009). Subjects in these experiments made paired lottery choices similar to that shown below:

> *You face the following pair of concurrent decisions. First examine both decisions, then indicate your choices, by circling the corresponding letter. Both choices will be payoff relevant, i.e., the gains and losses will be added to your overall payment.*
>
> **Decision (1)**. *Choose between:*
>
> *A. A sure gain of $2.40.*

B. A 25 percent chance to gain $10.00, and 75 percent chance to gain nothing $0.00

Decision (2). *Choose between:*

C. A sure loss of $7.50

D. A 75 percent chance to lose $10.00, and 25 percent chance to lose $0.00

Rabin and Weizsacker (2009)

28 percent of subjects in Rabin and Weizsacker (2009)[5] chose the pair of lotteries A and D, which, had only one of Decisions (1) and (2) been payoff-relevant, are consistent with prospect theory (Kahneman and Tversky, 1979) preferences. However both decisions *were* paid, and simple calculations show that the pair of choices AD induces exactly the same lottery as the pair BC, minus £0.10, indicating that the lottery induced by action pair AD is first-order stochastically dominated by action pair BC. The choice AD therefore cannot be rationalized with any expected utility model that assumes more money is strictly better than less. When subjects were presented the task as a choice between the four lotteries induced by the four possible action pairs, not one individual made this error, suggesting that in the first task, subjects either did not realize that there were benefits to bracketing broadly, or the costs of doing so were too high. Similar effects have been found when subjects evaluate investments over time. In these settings forcing subjects to consider a longer time period makes them more willing to take on risk (Gneezy and Potters, 1997; Thaler et al., 1997).

[5]Here I report the incentivized laboratory experiment of Rabin and Weizsacker (2009). Tversky and Kahneman (1981) and Rabin and Weizsacker (2009) also report much higher frequencies of this pair of dominated choices with hypothetical decisions.

Rabin and Weizsacker (2009, p1513) suggest a simple model for such behavior which can be taken to data. Players make I choices $\mathbf{m} = \{m_i\}_{i=1}^{I}$ to maximize:[6]

$$U(\mathbf{m}) = \kappa \int u(x^I) dF(x^I \mid \mathbf{m}) + (1-\kappa) \sum_i \int u(x_i) dL_i(x_i \mid m_i) \tag{1.1}$$

where u is a valuation for money, $F(\cdot \mid \mathbf{m})$ is the distribution of monetary outcomes conditional on choices $\mathbf{m}$, $L_i(\cdot \mid m_i)$ is the distribution of monetary outcomes from choosing action m_i for Decision i, and $\kappa \in [0,1]$ is the degree of broad-bracketing. That is, for $\kappa = 1$ the individual fully takes into account of the interactions between their decisions, and for $\kappa = 0$ the individual narrow brackets. This model therefore nests rational expected utility maximization ($\kappa = 1$) and narrow bracketing ($\kappa = 0$).

Choice bracketing is closely related to the treatment of background risk (Eeckhoudt et al., 1996). Under reasonably weak restrictions on risk preferences, a risk averse subject should be less willing to choose a risky option if his background risk worsens. Ignoring background risk in the Roommate's Dilemma is similar to narrowly bracketing because these subjects will ignore the risk they are exposed to in one game when making a decision in the other. In a field experiment, Harrison et al. (2007) demonstrate that subjects typically do pay attention to background risk, suggesting by this metric that broad bracketing should make better predictions in the Roommate's Dilemma.

Narrow bracketing is a more general case of mental accounting (Thaler, 1985). Individuals who mentally account assign artificial budgets to different types of purchases; for example, food and entertainment. If the price of (say) seeing a movie increases, this individual would substitute away from movies and toward other forms of entertainment. However their food consumption would not change, and so if movies and popcorn are substitutes, they could be made better off by substituting away from popcorn as well. Felsö and Soetevent (2014) identify shoppers who mentally account between cash and gift certificate wealth. They surveyed shoppers after spending a

[6] An alternative but similar specification is made by Barberis and Huang (2009).

gift certificate. Respondents typically indicated that while most (83%) would have bought the item had they not received the certificate, among the remaining 17%, 78% of these indicated that had they instead received a cash gift they would have spent the money on a everyday purchase. This suggests that the majority of shoppers who would have changed their behavior had they not received the gift certificate did so not because of restricted consumption possibilities, but because they bracketed cash and gift certificate wealth separately. While one typically thinks of mental accounting and choice bracketing as behaviors that make individuals (weakly) worse off, bracketing could help to mitigate other "behavioral" shortcomings. For example Koch and Nafziger (2016) shows theoretically that bracketing goals (such as weight loss) more narrowly can help overcome self-control problems. The main driver of this result is that loss-averse individuals cannot balance out losses and gains across brackets, so they are more motivated to meet these goals. Therefore, in addition to lowering the cost of decision-making, bracketing may be an evolutionary side-effect of self-control problems.

While studies on individual choice behavior have documented this behavior well, it is not directly clear that such behavior will also be present in strategic interactions. Once opponents' strategies are fixed, any game is reduced to an individual choice problem in which the individual could be a narrow bracketer. Players may be narrow or broad bracketeres, and may have beliefs about the bracketing ability of their opponents. That is, individuals may play games differently to predictions made by theories assuming only broad bracketers for two reasons: (1) they may narrowly bracket, and therefore fail to recognize dominated action pairs as in the lottery choice problem described above, or (2) they may believe that their opponents will narrowly bracket, and therefore adjust their actions based on the behavior they expect their opponents to take.

One study that could be informative on choice bracketing behavior in strategic settings is Blanco et al. (2010). These authors analyzed whether subjects change their behavior when there is a hedging opportunity available due to a paid beliefs elicitation task. Choice bracketing is important here because a narrow bracketer would fail to recognize that by stating incorrect beliefs one could increase the minimum possible earnings from the two tasks. In a sequential prisoner's dilemma, subjects behaved no differently when there were hedging opportunities than when there were not, supporting narrow bracketing. On the other hand in a coordination game, when hedging opportunities were more obvious, players not only recognized and used the hedging opportunity, but post-experiment surveys revealed that subjects adjusted their play in the coordination game because they thought others would recognize hedging opportunities. These results suggest that players may narrow bracket unless the benefits from broad bracketing are obvious to them, and that observing play consistent with narrow bracketing (and not consistent with broad bracketing) may be better explained by whether subjects are able to recognize that benefits of bracketing broadly exist, rather than that subjects' mental costs of bracketing broadly are prohibitively high.

This paper is also somewhat related to the literature on game-form recognition, which studies the effect of displaying games in different formats. McCabe et al. (2000) observe more cooperation when a sequential-move game is presented in extensive form compared to when it is presented in strategic form; Chou et al. (2009) demonstrate that deviations from game-theoretic predictions in the 2-person beauty contest game (namely choosing weakly dominated actions) can be remedied by presenting the game with a hint about the best response; in a non-strategic setting, Cason and Plott (2014) demonstrate that the accuracy of elicited preferences is harmed by subjects' understanding of their task, but can be remedied through repetition with feedback. Finally Cox and James (2012) study centipede games and Dutch auctions presented in clock (standard for Dutch auction experiments) and tree (standard for centipede game experiments) formats, concluding surprisingly that behavior is closer to equilibrium predictions when the non-standard game-form representation is used.[7]

[7]It should be noted here that game theory's treatment of the clock and tree format changes, somewhat, the definition of a strategy. Strategies in the tree format must specify an action even for histories that the strategy rules out. The tree/clock treatment effect therefore may not be entirely due to game form recognition.

In all but one of these cases, we observe treatment effects in strategic settings that are not predicted by standard theory. If the representation effects in individual choice experiments of Tversky and Kahneman (1981) and Rabin and Weizsacker (2009) are also relevant in strategic setting, then playing multiple, payoff-relevant games at the same time may have different outcomes depending on whether individuals are able to integrate the lotteries induced by opponents' strategies in the two games to fully understand the distribution of payoffs, and that presenting the "games" in Table 1.1 separately may result in treatment effects inconsistent with standard models of rationality used in economics. While the current experimental design does not vary the game form, one could hypothesize that there may be different representations of the Roommate's Dilemma which either reduce or increase the cognitive cost of decision-making, potentially altering the results.

1.2 Experiment design

The purpose of this experiment is to determine whether narrow bracketing plays a role in strategic settings, and whether narrow bracketing in non-strategic settings is a predictor for narrow bracketing in strategic settings. To this end, I use a 3-treatment between-subjects design, varying the payoffs systematically between each treatment. These treatments are paired with a lottery choice task that aims to measure players' risk preferences, as well as whether they narrowly or broadly bracket in the absence of strategic considerations.

In the following discussion, I state the hypotheses as the intended null, and the expected result of the test under the relevant assumption (i.e. either broad or narrow bracketing). All hypotheses are designed to be testable with simple comparisons of means between treatments, however logistic regressions are also used to provide further insight and control for more factors.

Table 1.2.: Treatments for the Roommate's Dilemma. Payoffs are the sum of outcomes in both tables.

Treatment 1

Γ_1

	A	B
A	10 10	100 10
B	10 100	0 0

Γ_2

	C	D
C	56 56	160 56
D	56 160	0 0

Treatment 2

Add 50 points to Γ_1 in Treatment 1

Γ_1

	A	B
A	60 60	150 60
B	60 150	50 50

Γ_2

	C	D
C	56 56	160 56
D	56 160	0 0

Treatment 3

Add 50 points to Γ_2 in Treatment 1

Γ_1

	A	B
A	10 10	100 10
B	10 100	0 0

Γ_2

	C	D
C	106 106	210 106
D	106 210	50 50

1.2.1 The Roommate's Dilemma

The purpose of this part of the experiment is to test comparative static predictions based on assumptions of broad and narrow bracketing. The three treatments used to test these assumptions are shown in Table 1.2. Note that compared to the Treatment 1, Treatment 2 adds 50 points to Γ_1, and Treatment 3 adds 50 points to Γ_2. All other aspects of the games are identical. In all treatments, any mixed-strategy Nash equilibrium assuming risk neutrality satisfies $(p_A, p_C) = (0.10, 0.35)$, however varying payoff levels in this way will affect players' incentives of they are not risk neutral. I therefore proceed under the working assumption that players are not risk neutral, and derive some comparative static predictions assuming that they are risk averse. This assumption can be tested either by comparing choice probabilities between all three treatments: under risk neutrality they should all be equal, irrespective of whether subjects broadly or narrowly bracket. Additionally, individual risk preferences are elicited in the Lottery Task.

In order to test an assumption of broad bracketing, we can compare Treatments 2 and 3. This is because the payoffs of the full, 2-player 4-action game are identical. A broad bracketer, having combined the payoffs of Γ_1 and Γ_2, would not distinguish between Treatments 2 and 3 (in particular, she could see these treatments both as "Treatment 1 plus 50 points"). Therefore, if subjects broad bracket in games, we can expect to not reject:

Hypothesis 1 (Broad bracketing) *Play in Treatments 2 and 3 are identical.*

Narrow bracketers consider Treatments 2 and 3 to be different because they do not add up the payoffs between Γ_1 and Γ_2. However for both Γ_1 and Γ_2 in Treatment 1, there exists an identical game in either Treatment 2 or 3. In particular: Γ_1 is the same in Treatments 1 and 3, and Γ_2 is the same in Treatments 1 and 2. As narrow bracketers will ignore one game entirely while making their decision in the other, choice probabilities within a (bracketed) game will not change unless the payoffs *in that game* change. Therefore, if subjects narrowly bracket in games, we can expect not to reject:

Hypothesis 2 (Narrow bracketing, weak version) *Play in* Γ_1 *(i.e.* p_A*) will be identical in Treatments 1 and 3, and play in* Γ_2 *(i.e.* p_C*) will be identical in Treatments 1 and 2.*

Furthermore, placing additional structure on risk preferences, one can make comparative static predictions based on narrow bracketing in the games where the payoffs change. For example, by assuming decreasing absolute risk aversion (DARA), the following result can be used to show that adding points to a game will decrease the equilibrium probability of taking the safe action in the relevant game (i.e. actions A and C):

Theorem 1.2.1 (Eeckhoudt et al. (1996), p685, Proposition 1) *Suppose that the individual has decreasing absolute risk aversion. A FSD [first-degree stochastic dominance] deterioration in background risk would make the individual uniformly more risk-averse if it takes the form of adding a negative noise that is statistically independent to the initial background risk:* $\tilde{y}_2 =_d \tilde{y}_1 + \tilde{\epsilon}$, *with* $\Pr[\tilde{\epsilon} \leq 0] = 1$ *and* $\tilde{\epsilon}$ *and* $\tilde{y}_1$ *are independently distributed.*

Therefore assuming DARA, a wide class of risk preferences, and narrow bracketing, we can expect to not reject:

Hypothesis 3 (Narrow bracketing, strong version) *Play in* Γ_1 *(i.e.* p_A*) will be identical in Treatments 1 and 3, and play in* Γ_2 *(i.e.* p_C*) will be identical in Treatments 1 and 2. Compared to Treatment 1,* p_A *will be higher in Treatment 2, and* p_C *will be higher in Treatment 3.*

That is, broad bracketing and DARA predicts that compared to Treatment 1, p_A and p_C will be lower in Treatment 2 and 3 respectively. For example, for narrow bracketing and CRRA preferences (a special case of DARA), the mixed-strategy Nash equilibrium in Γ_1, where δ_1 points are added to this game, is characterized by:

$$p_A(\delta_1) = \frac{(10 + \delta_1)^{1-r} - \delta_1^{1-r}}{(100 + \delta_1)^{1-r} - \delta_1^{1-r}} \tag{1.2}$$

which is strictly decreasing in δ_1: the utility function is less concave at higher levels of wealth, so agents are more willing to take on risk as their baseline wealth increases.

Finally, it should be noted that this design permits tests of the competing theories not against each other, but against "something else". Rejection of one of Hypotheses 1-3 therefore is not support for the other model. While some may see this as a weakness of the design, the author believes that this provides a more level playing field for the competing theories. The design gives both an opportunity to succeed or fail.

1.2.2 Lottery task

The purpose of the lottery task is to estimate subjects' risk aversion and bracketing behavior. In particular, for every subject I aim to elicit parameter κ (bracketing behavior) from Equation 1.1, and the constant relative risk aversion parameter r_R. This is achieved through a convex budget set problem, where the subject must decide how many tokens to assign to a risky investment. The risky investment pays out $\chi > 2$ times the number of tokens invested with probability 50%, and zero otherwise. In order to pin down these parameters more accurately, subjects make decisions for several different values of χ.

The lottery task is divided into two parts. In the first, the subject is asked to make choices for just one payoff relevant investment. In the second, there are two payoff-relevant investments. Therefore behavior in the first task is not affected by bracketing behavior, while in the second it is. In total, subjects made four decisions in Part 1 of the Lottery Task, and four decisions in Part 2. One of these eight decisions was randomly chosen for payment. This avoids confounds associated with subjects broadly bracketing across different instances of the task.

Part 1: Elicitation of risk preferences in the absence of bracketing considerations

In this task, subjects are endowed with 100 tokens. They can invest any number of these in the risky asset, which returns $\chi > 2$ times the number of tokens invested with probability 50%, otherwise asset returns zero. Subjects play four instances of this task for $\chi = 2.25,\ 2.75,\ 3.50,\ 5.50$, which were calibrated so that an individual would invest half of their tokens if they had CRRA preferences of approximately[8] $r = 0.2,\ 0.4,\ 0.6,\ 0.8$ respectively. That there is only one payoff-relevant investment decision ensures that narrow and broad bracketers with the same risk preferences behave identically in this part.

[8] χ is rounded off to the nearest 0.25.

Ignoring the indivisibility of the tokens and normalizing the mass of tokens to 1, a subject's utility maximization problem in this task is:

$$\max_{x\in[0,1]} U(x) = 0.5u(\chi x + 1 - x) + 0.5u(1 - x) \tag{1.3}$$

where x is the fraction invested in the risky asset.

The first-order condition is:

$$0 = (\chi - 1)u'(\chi x^* + 1 - x^*) - u'(1 - x^*) \tag{1.4}$$

For CRRA preferences, this reduces to:

$$\begin{aligned} 0 &= (\chi - 1)(\chi x^* + 1 - x^*)^{-r} - (1 - x^*)^{-r} \\ -r\log(1 - x^*) &= \log(\chi - 1) - r\log(\chi x^* + 1 - x^*) \\ r &= \frac{\log(\chi - 1)}{\log(\chi x^* + 1 - x^*) - \log(1 - x^*)} \end{aligned}$$

x^* is strictly decreasing in r and strictly increasing in χ.

Part 2: elicitation of bracketing behavior

Subjects' task in Part 2 differs from Part 1 only in that in each instance there are two investment decisions to be made. In this part subjects have 100 red tokens, and 100 blue tokens. The red investment returns $\chi > 2$ with probability 50%, zero otherwise, and the blue investment returns $\chi > 2$ with probability 50%, zero otherwise. The success of these investments are independent. Both red and blue tokens count toward a subject's earnings.

Subjects play four instances of this task, with $\chi = 2.20,\ 2.50,\ 3.00,\ 4.00$. As in Part 1, these values were chosen so that a broad bracketer with CRRA preferences $r = 0.2,\ 0.4,\ 0.6,\ 0.8$ respectively would invest approximately 70% of her tokens in each asset.[9] Assuming DARA preferences, by Theorem 1.2.1 a narrow bracketer will invest less than a broad bracketer who is equally risk averse. This is used to separate broad and narrow bracketers. Targeting investments for the broad bracketers in a reasonably higher region of the allowable range compared to Part 1 means that there is substantial room for the narrow bracketers to reveal themselves before running into the corner solution of not investing anything.

Predictions and simulation results

Figure 1.1 shows predictions in the Lottery Task for CRRA individuals for broad bracketers ($\kappa = 1$) and narrow bracketers ($\kappa = 0$). These predictions are consistent with the general predictions that (1) individuals who are more risk averse invest less, and (2) individuals who bracket more invest more in Part 2 (i.e. Theorem 1.2.1).

Of particular interest in Figure 1.1 is the identification power for r and κ. For r, performing Task 1 for only $\chi = 2.25$ would, in principle, be sufficient to identify risk aversion for all but the most risk-neutral individuals; however the four instances of Task 1 provide a robustness check and will help mitigate the effect of noisy behavior on estimation. For κ, by having several instances of Task 2, we increase the chance of narrow and broad-bracketing predictions being widely different in at least one instance for each type. For example, the blue lines in Figure 1.1 show the difference between broad and narrow choices in Task 2. For $\chi = 2.2$ there is a large difference for low r, and for the other cases the difference is large later. This means that for each r there is good separation of bracketing behavior in at least one case. While this statement does not apply for individuals who are very close to risk neutral, I am not as concerned with estimating κ for r this low because the benefits of bracketing broadly over bracketing narrowly for these individuals are small, and zero for risk-neutral subjects.

[9] Here I choose the higher target of 70% (compared to 50% in Part 1) because narrow bracketers will invest less, thus ensuring a large variability of investment with bracketing behavior.

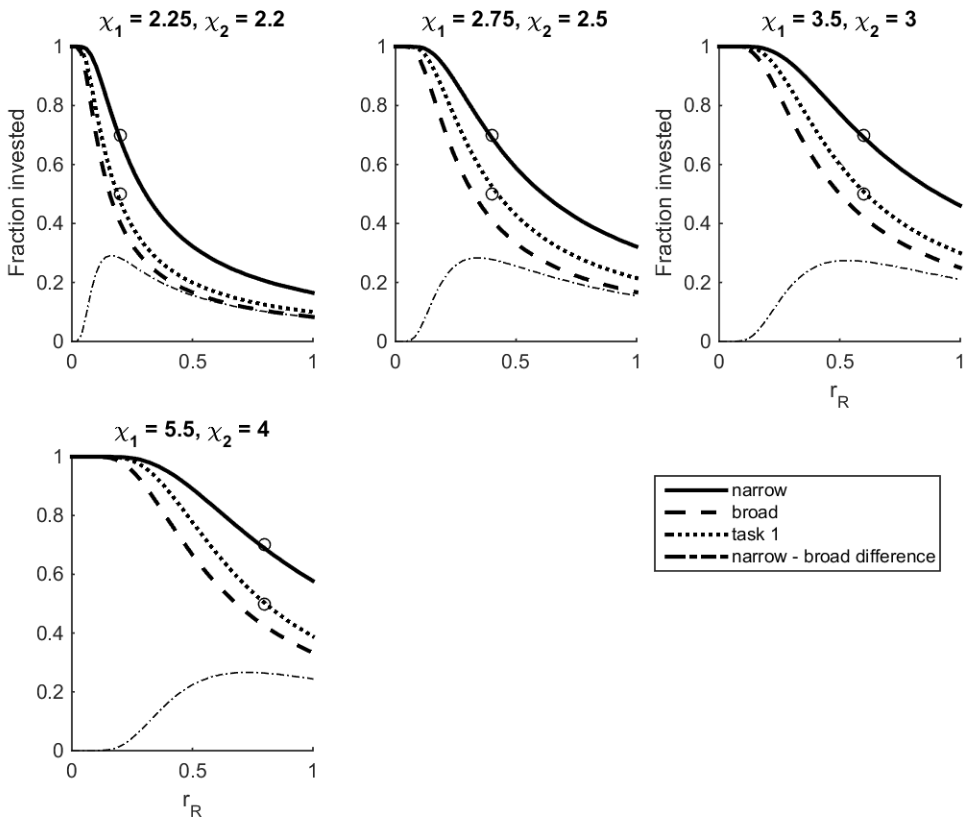


Fig. 1.1. Predictions for behavior in the lottery task for subjects with CRRA preference r and bracketing behavior. The dotted curve shows predicted behavior in Part 1, which does not change with bracketing behavior. The solid and dashed lines show predictions for Part 2 for broad bracketing and narrow bracketing respectively. Circles show the targeted behavior for each case of the task, and the dash-dotted lines show the difference between broad and narrow bracketing choices. For all targeted preferences, this difference is a significant fraction of the action space for at least one instance of Part 2, which makes separation of narrow and broad bracketers possible for a wide range of risk preferences observed in the lab.

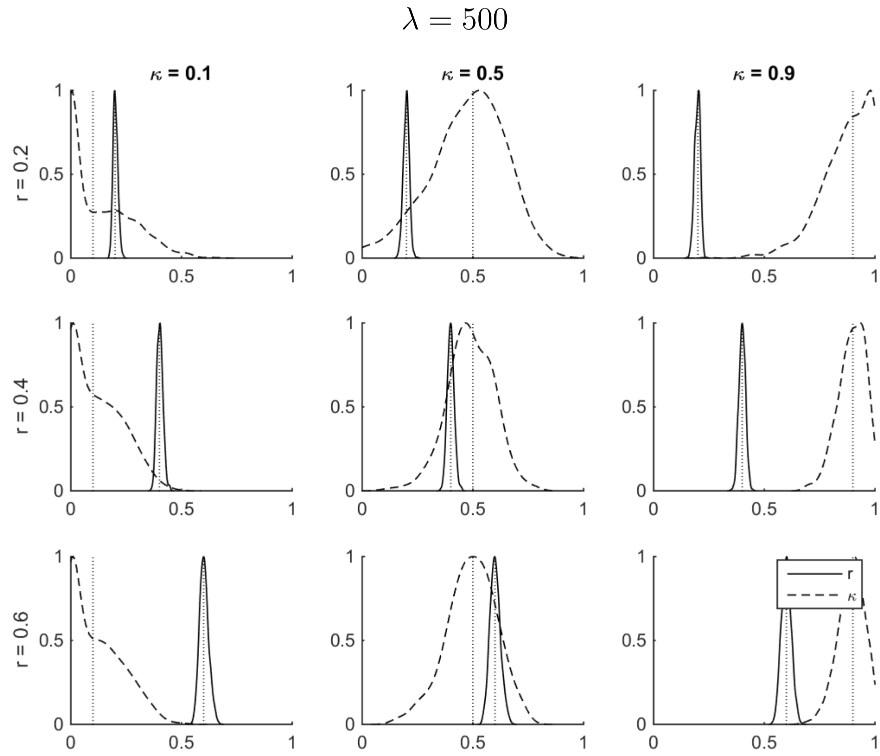


Fig. 1.2. Simulation results for maximum likelihood estimation of parameters r and κ. Choices simulated with Logit precision parameter $\lambda = 500$ and the investment space restricted to multiples of 2% of the total tokens. The mass of tokes is normalized to 1. Row and column labels show the true values of the parameters being estimated.

The properties of the lottery task are generally born out in the simulation shown in Figure 1.2, where estimates of r are generally very close to the true value, and estimates of κ are more precise as r increases.

1.2.3 Implementation

The experiment consists of three parts: the Roommates' Dilemma, followed by the two lottery tasks. Before the beginning of each part, subjects are given hard copies of their instructions (see Appendix A.4). These instructions are read aloud by an experimenter. Following this, subjects complete a questionnaire on their computers to test their understanding of the instructions. In lieu of a show-up fee, subjects are paid 20 experimental dollars (US$0.40) for each correctly answered question in these questionnaires, irrespective of whether that part is randomly chosen for payment. There were 12 such questions in total.

Subjects were recruited from the undergraduate population of Purdue University, West Lafayette, who had registered for economics experiments using ORSEE (Greiner, 2015). The experiment took place in the Vernon Smith Experimental Economics Laboratory at Purdue University between November 2014 and April 2015, and was programmed using z-Tree (Fischbacher, 2007). Overall, 128 subjects participated in the experiment.

At the end of the experiment, subjects received feedback on all decisions that were chosen for payment, were paid privately in cash, completed an unpaid survey, and were allowed to leave the laboratory without interacting with other participants.

Part 1 – Roommates' Dilemma: At the beginning Part 1, subjects were randomly assigned into groups of four, which remained fixed for the rest of Part 1. Subjects played exactly one treatment of the Roommates' dilemma for 20 periods. In each period, subjects were randomly matched with an opponent from their group, and choose a pair of actions (i.e. one of $A\&C$, $A\&D$, $B\&C$, and $B\&D$) for that period. In every period except the first, subjects saw a history table displaying their choices, the choices of their randomly chosen opponent, and the frequency of each action taken by all all four group members. Screenshots of all parts are shown in Appendix A.5. The random matching encourages subjects to think about mixed strategies because it is at best somewhat difficult to coordinate on pure strategies when one's partner is changing randomly every period. As subjects interacted within a group of four, each group is an independent observation for the Roommate's Dilemma, and analysis in later sections of the paper will either cluster or aggregate to appropriately respect this.

In order to minimize hedging opportunities (outside of those available to broad bracketers) while still making choices payoff-relevant, five of the twenty periods were randomly chosen for payment at the end of the experiment.

Parts 2 & 3 – Lottery task: Subjects completed parts 1 and 2 of the lottery task after completing the Roommates' Dilemma part. One decision from the eight made in these parts was randomly chosen for payment. If the randomly chosen decision was from Part 2 of the lottery task, this was for both the red and the blue lottery.

Payments: As described above, subjects made six payoff-relevant decisions, plus the earnings from the questionnaires. Experimental dollars were converted to US dollars at a rate of 1 experimental dollar = US$0.02. In Treatment 1, payoffs could range from US$0.00 to US$46.80. Treatments 2 and 3 add 50 points to five payoff-relevant decisions, so payments in these treatments could range between US$5.00 and US$51.80. All but (at most) US$4.80 from the questionnaire was earned through decisions in the experiment. At the time of running the sessions, both the US and Indiana hourly minimum wages were US$7.25. Sessions typically lasted about 75 minutes.

1.3 Results

1.3.1 Aggregate play in the Roommate's Dilemma

Result 1 *Aggregate behavior in the Roommate's Dilemma is not consistent with broad bracketing.*

Support: Figure 1.3 shows the aggregate choice probabilities between treatments of all four action pairs. If broad bracketing were a good predictor of play, then in line with Hypothesis 1 the gray bars (Treatment 2) would be the same height as the white bars Treatment 3) While there is little treatment effect in the probabilities of choosing very safe (AC) or very risky (BD) actions, the "in between" actions AD

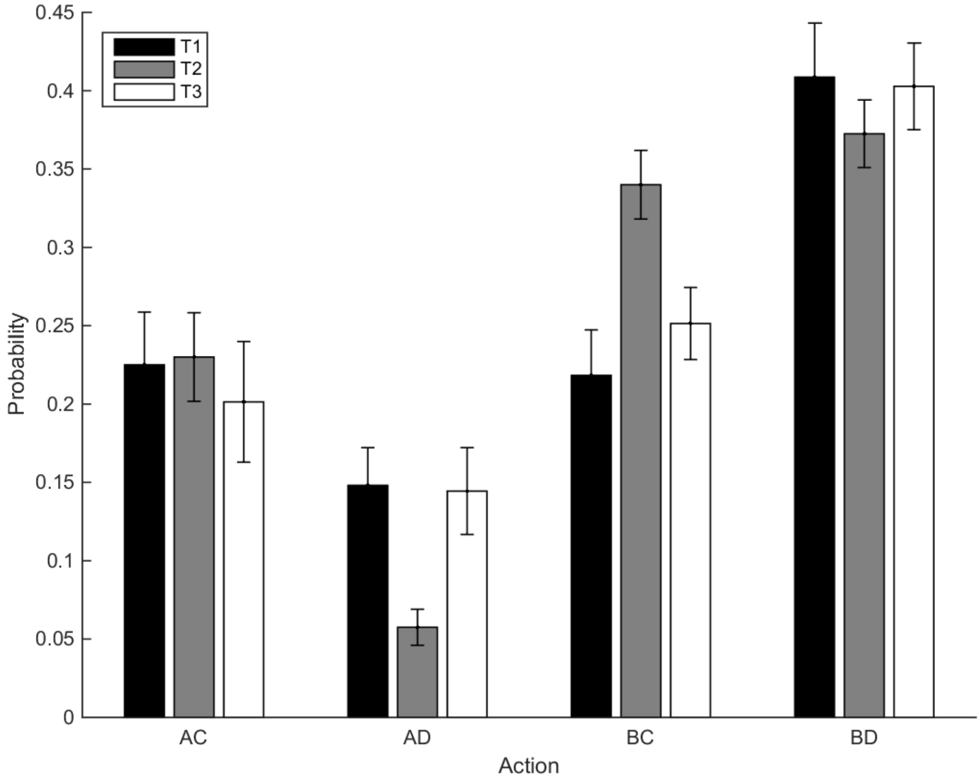


Fig. 1.3. Choice probabilities of all four action pairs in the Roommate's dilemma by treatment. Whiskers show ± one standard error. If play is predicted well by broad bracketing, the gray and white bars (Treatments 2 and 3 respectively) would be the same heights.

Table 1.3.: Permutation tests that choice probabilities from pairs of treatments are equal. The first two rows show tests that mean group choice probabilities in Γ_1 and Γ_2 are the same respectively, the middle four rows test combinations of actions over the two games, again using mean group choice probabilities. The final two rows test that the joint choice probabilities are equal. "Narrow" tests jointly that p_A and p_C are equal across treatments, and "Broad" tests that p_{AC}, p_{AD}, p_{BC}, and p_{BD} are equal across treatments. Except for the last two rows, the test statistic is $\frac{\hat{\mu}_1 - \hat{\mu}_2}{\sqrt{\hat{\sigma}_1^2/n_1 + \hat{\sigma}_2^2/n_2}}$, the standard t-statistic for testing equal means. The final two rows report a likelihood ratio statistic, which would be asymptotically distributed χ_3^2 under the null (the asymptotic assumption is not required for the permutation test). 2-sided p-values in parentheses are computed by permuting the samples 1000 times. The sample is restricted to the final 10 periods for this analysis.

Action	1 & 2	Action	1 & 3	Action	2 & 3	
A	1.418	(0.191)	0.618	(0.548)	-0.878	(0.378)
C	-3.897	(0.000)	-1.685	(0.122)	2.280	(0.046)
AC	-0.424	(0.712)	0.914	(0.383)	1.422	(0.170)
AD	2.941	(0.010)	-0.125	(0.877)	-3.210	(0.003)
BC	-2.647	(0.011)	-2.322	(0.047)	0.774	(0.454)
BD	0.897	(0.373)	1.261	(0.229)	0.458	(0.609)
Narrow (LR)	44.089	(0.000)	1.526	(0.479)	26.860	(0.000)
Broad (LR)	62.638	(0.000)	3.181	(0.373)	41.958	(0.000)

and BC are noticeably different. Specifically, subjects tend to play BC more in Treatment 3, and AD more in Treatment 2. Table 1.3 shows tests for equal means of a particular choice probability in two treatments. Of interest to Result 1 are the tests reported in the rightmost column: the comparison of actions in Treatments 2 and 3. Broad bracketing predicts that there should be no significant difference between these treatments, yet we see that action AD (as well as C) is played significantly more frequently in Treatment 3 (two-sided $p = 0.003$).

Further evidence against broad bracketing in games can be found in Table 1.4, which shows the results of multinomial logistic regressions of actions in the Roommates Dilemma, restricting analysis to Treatments 2 and 3. Of particular interest are the bottom two rows of this table, which report a test of Hypothesis 1. In all but Model (1), we reject this hypothesis at reasonable significance levels.[10] It should be

[10]The failure to reject in Model (1) is unsurprising as the sample is restricted to actions in the first period only, where play is most likely to be noisy, and the power of this test is low.

Table 1.4.: Multinomial logistic regression of Roommate's Dilemma actions in Treatments 2 and 3, controlling for period effects in various different ways. *AC* is the base outcome.

	(1)	(2)	(3)	(4)	(5)
AD					
Treatment==2	-1.387***	-1.054**	-1.058**	-1.045**	-1.063**
	(0.375)	(0.332)	(0.334)	(0.360)	(0.337)
Period^{-1}	8.116		-1.174**	-1.145*	
	(16.83)		(0.398)	(0.468)	
$\text{Period}^{-1}\times$ (Treatment==2)				-0.0480	
				(1.005)	
Constant	-0.575	-0.332	-0.0947	-0.102	-1.132*
	(1.057)	(0.286)	(0.277)	(0.258)	(0.495)
BC					
Treatment==2	-0.224	0.169	0.165	0.0706	0.172
	(0.272)	(0.289)	(0.294)	(0.285)	(0.299)
Period^{-1}	13.04		-1.147***	-1.533***	
	(12.30)		(0.268)	(0.422)	
$\text{Period}^{-1}\times$ (Treatment==2)				0.591	
				(0.543)	
Constant	-0.214	0.222	0.455	0.515*	-0.374
	(0.804)	(0.238)	(0.238)	(0.219)	(0.346)
BD					
Treatment==2	-0.266	-0.211	-0.215	-0.0892	-0.212
	(0.262)	(0.266)	(0.272)	(0.297)	(0.276)
Period^{-1}	6.247		-1.559***	-1.210**	
	(9.082)		(0.344)	(0.470)	
$\text{Period}^{-1}\times$ (Treatment==2)				-0.777	
				(0.696)	
Constant	0.563	0.693**	0.994***	0.934***	-0.312
	(0.578)	(0.213)	(0.224)	(0.231)	(0.336)
Observations	760	1520	1520	1520	1520
Cluster level	Group	Group	Group	Group	Group
Number of clusters	19	19	19	19	19
Period restriction	11-20	none	none	none	none
Period dummies	N	N	N	N	Y
$H0: T2 = T3, \chi^2(3)$	14	27.1	27.2	11.2	27.3
p-value	.0029	0	0	.0107	0

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

noted, however, that Hypothesis 1 is rejected not in favor of narrow bracketing, but in favor of *anything else*: rejecting Hypothesis 1 here is not support for narrow bracketing, only support against broad bracketing. We now turn to testing the predictions of narrow bracketing.

Result 2 *Aggregate behavior in the Roommate's Dilemma is not consistent with narrow bracketing.*

Support: Figure 1.4 shows the choice probabilities in the Roommate's Dilemma, marginalized over Γ_1 and Γ_2. According to Hypothesis 2, narrow bracketing predicts that p_A should not change between Treatments 1 and 3, and p_C should not change between Treatments 1 and 2. Inspecting this figure, the latter is not supported by the data. This is supported in Table 1.3, which indicates that C is played more frequently in Treatment 2, compared to Treatment 1 (2-sided $p < 0.001$). This is not consistent with Hypothesis 2. Additionally, note that Hypothesis 3 is not supported as there is no significant difference between Treatments 1 and 2 in the probability of choosing A, and no significant difference between Treatments 1 and 3 in the probability of choosing C. In fact, if one also assumes that subjects are risk-averse, the probability of choosing C moves in the wrong direction between Treatments 1 and 3.

These results are also supported by Table 1.5, which reports estimates from a bivariate probit regression using the actions A and C as dependent variables. In particular, the only action whose whose probability varies by treatment is action C, which in Treatment 2 is significantly different from its base case choice probability in Treatment 1. This pattern of changes, or lack thereof, between treatments is support against both hypotheses made assuming that subject bracket narrowly. Firstly, Hypothesis 2 can be rejected because the Treatment 2 dummy on action C is significantly different from zero. The joint test of *all* predictions made in Hypothesis 2 , namely that p_A does not change between Treatments 1 and 3, and p_C does not change between Treatments 1 and 3, is reported in the final two rows of the table. For all specifications we reject jointly that these probabilities are different. Secondly,

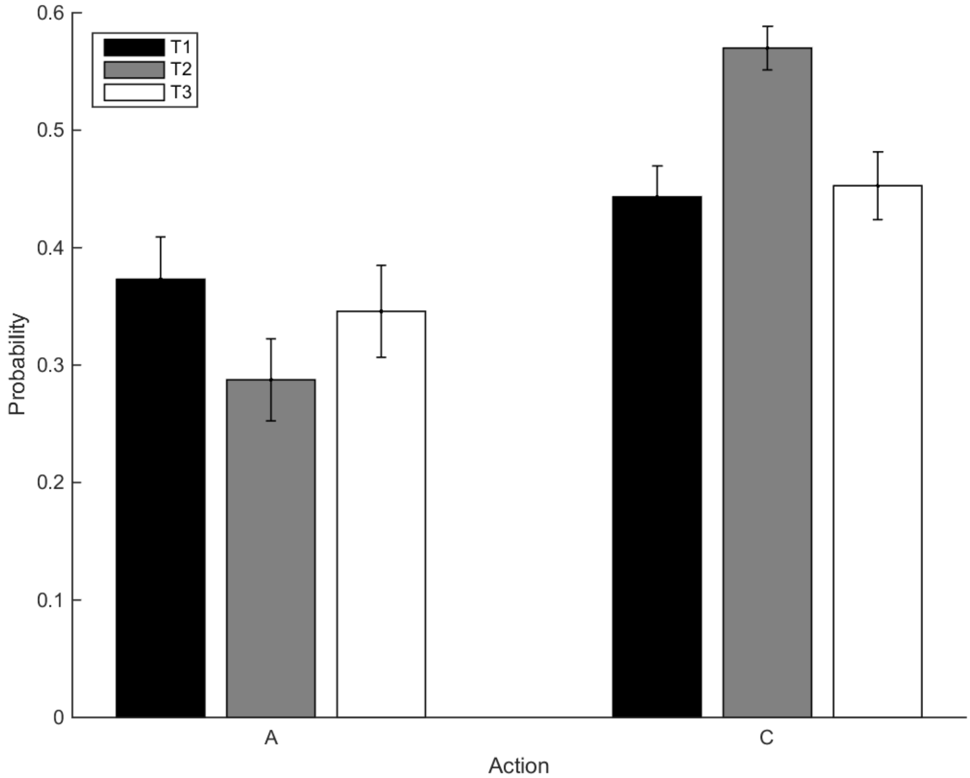


Fig. 1.4. Choice probabilities by treatment in the Roommate's Dilemma, marginalized over Γ_1 and Γ_2. Whiskers show $\pm$ one standard error.

Table 1.5.: Bivariate probit regression of actions in Γ_1 and Γ_2.

	(1)	(2)	(3)	(4)	(5)
A					
Treatment==2	-0.221	-0.242	-0.245	-0.236	-0.249
	(0.147)	(0.130)	(0.132)	(0.149)	(0.133)
Treatment==3	-0.0779	-0.0704	-0.0705	-0.0275	-0.0715
	(0.135)	(0.130)	(0.132)	(0.146)	(0.133)
Period^{-1}	-0.780		0.722^{***}	0.806^{***}	
	(2.896)		(0.132)	(0.217)	
$\text{Period}^{-1}\times$ (Treatment==2)				-0.0519	
				(0.316)	
$\text{Period}^{-1}\times$ (Treatment==3)				-0.238	
				(0.319)	
Constant	-0.386	-0.324^{***}	-0.455^{***}	-0.470^{***}	0.117
	(0.232)	(0.0893)	(0.0929)	(0.0967)	(0.139)
C					
Treatment==2	0.335^{***}	0.321^{***}	0.324^{***}	0.371^{***}	0.326^{***}
	(0.0838)	(0.0763)	(0.0778)	(0.0773)	(0.0780)
Treatment==3	0.131	0.0243	0.0245	0.159	0.0246
	(0.0743)	(0.0904)	(0.0917)	(0.0898)	(0.0920)
Period^{-1}	0.767		0.770^{***}	1.080^{***}	
	(2.430)		(0.120)	(0.172)	
$\text{Period}^{-1}\times$ (Treatment==2)				-0.266	
				(0.235)	
$\text{Period}^{-1}\times$ (Treatment==3)				-0.763^{**}	
				(0.276)	
Constant	-0.315	-0.143^{*}	-0.280^{***}	-0.334^{***}	0.414^{***}
	(0.168)	(0.0626)	(0.0652)	(0.0653)	(0.107)
athrho					
Constant	0.387^{***}	0.416^{***}	0.397^{***}	0.396^{***}	0.401^{***}
	(0.0897)	(0.0738)	(0.0736)	(0.0739)	(0.0735)
Observations	1280	2560	2560	2560	2560
Cluster level	Group	Group	Group	Group	Group
Number of clusters	32	32	32	32	32
Period restriction	11-20	none	none	none	none
Period dummies	N	N	N	N	Y
H0 : Narrow, $\chi^2(2)$	17.2	20.7	20.3	24.3	20.3
p-value	.0002	0	0	0	0

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Hypothesis 3 is not supported since we fail to reject significant changes in either p_A moving from Treatment 1 to 2, or in p_C moving from Treatments 1 to 3.

1.3.2 Structural tests of behavior

While Section 1.3.1 tested the predictions of narrow and broad bracketing by comparing treatment effects, this section pools data from all treatments and aims to explain decision-making in both tasks on the individual level. Instead of forcing behavioral parameters (such as those describing risk aversion and bracketing) to be constant across the two tasks, I allow them to be different but possibly correlated. Estimating a positive correlation between bracketing in the two tasks can therefore be interpreted as subjects either bracketing broadly or narrowly in each task. I assume that the behavior of subjects is characterized by five individual-level parameters:

r_i^{RD}, r_i^{LT} describe risk aversion in the Roommate's Dilemma and Lottery Task respectively. Here I assume the utility function: $u_i(x) = x^{r_i}, \quad r_i \in \mathbb{R}_{++}$.[11]

$\kappa_i^{RD}, \kappa_i^{LT}$ describe subject i's bracketing behavior in the Roommate's Dilemma and the Lottery Task respectively, according to (1.1). These are binary variables taking on values of 0 and 1.

γ_i models subject i's learning about her opponents' strategies during the Roommate's dilemma. I use the model of Cheung and Friedman (1997) to describe how subjects update beleifs about opponents' strategies between rounds of the Roommate's Dilemma. Given a sequence of opponents' actions leading up to period t $\{s_{a,u}\}_{u=1}^t$, subject i's beliefs that action a will be taken in period $t+1$ is:[12]

$$\hat{s}_{a,t+1} = \frac{s_{a,t} + \sum_{u=1}^{t-1} \gamma_i^u s_{a,t-u}}{1 + \sum_{u=1}^{t-1} \gamma_i^u} = \frac{\sum_{u=0}^{t-1} \gamma_i^u s_{a,t-u}}{\sum_{u=0}^{t-1} \gamma_i^u} \tag{1.5}$$

For example, in period 16 subject i believes that her opponent will play action AD in the next period with probability $\hat{s}_{AD,17}$. As in Cason et al. (2010), I

[11] Note that this is not the CRRA function used in the examples in previous sections, but is commonly used in structural estimation of experimental data on risk aversion (see for example Harrison and Rutström, 2009)

[12] See Equation 2.1 on p49 of Cheung and Friedman (1997).

assume that $\gamma_i \in [0,1]$, with $\gamma_i = 0$ corresponding to naïve best reply to actions in the previous period, and $\gamma_i = 1$ representing fictitious play beliefs.

To account for individual-level heterogeneity and the parameter restrictions, I assume that these parameters are drawn from the underlying population distribution:

$$\begin{bmatrix} \log(r_i^{RD}) \\ \log(r_i^{LT}) \\ \Phi^{-1}(\gamma_i) \\ \kappa_i^{*RD} \\ \kappa_i^{*LT} \end{bmatrix} \sim \mathcal{MVN}\left(X_i\beta, \Sigma\right), \quad \kappa_i^j = I(\kappa_i^{*j} \geq 0),\ j \in \{RD, LT\} \tag{1.6}$$

where $\Phi(\cdot)$ is the standard normal cdf, $I(\cdot)$ is the indicator function, X is a $N \times S$ vector of subject characteristics, β is a $S \times 5$ matrix of mean coefficients, and Σ is a 5×5 covariance matrix. I restrict Σ so that the risk-aversion and beliefs parameters are uncorrelated with the latent bracketing parameters:

$$\Sigma_{i,j} = 0 \text{ if } i \in \{1,2,3\} \text{ and } j \in \{4,5\} \tag{1.7}$$

$$\Sigma_{i,j} = 0 \text{ if } i \in \{4,5\} \text{ and } j \in \{1,2,3\} \tag{1.8}$$

I further replace an identifying restriction on the variances of κ_i^{*RD} and κ_i^{*LT} so that they equal 1. Columns 4 and 5 of β can therefore be interpreted as the coefficients of a probit model for the marginal distributions of the bracketing parameters. Importantly, $\Sigma_{4,5} = \Sigma_{5,4} \in (-1,1)$ is the correlation between the latent variables. If bracketing behavior does not vary across tasks, then we would expect $\Sigma_{4,5}$ to be close to 1.

For both tasks, I adopt a logistic choice rule, which for any two actions a, a' in the action space, the probabilities $p(a)$ and $p(a')$ that these actions are taken satisfies:

$$\log(p(a)) - \log(p(a')) = \lambda[u(a) - u(a')] \tag{1.9}$$

where $\lambda > 0$ is the choice precision. I allow the precision to vary by task, but not by subject. These parameters are λ^{RD} and λ^{LT}.

I use a Bayesian approach to estimate the population parameters $(\beta, \Sigma, \lambda^{RD}, \lambda^{LT})$. Through the data augmentation process, I also obtain shrinkage estimates of the individual parameters $(r_i^{RD}, r_i^{LT}, \gamma_i, \rho_i^{RD}, \rho_i^{LT})$, where $(\rho_i^{RD}, \rho_i^{LT})$ are the probabilities that individual i brackets broadly in each task. Details of this procedure are outlined in Appendix A.3.

Result 3 *The majority of subjects narrowly bracket in the Roommate's Dilemma. The majority of subjects narrowly bracket in the Lottery task. Broadly bracketing in one task does not predict broadly bracketing in the other task.*

Support: Table 1.6 shows posterior moments of the estimated model. This model estimates that approximately 17% of subjects broadly bracket in the Roommate's Dilemma, and 5% in the Lottery Task. These fractions are marginally different to each other: the posterior probability that a subject is more likely to bracket broadly in the game compared to the lottery task is 93.63% The correlation between the latent variables for bracketing is not significantly different from zero at the 5% level of significance. Analysis at the individual level reveals a similar story. Figure 1.5 shows shrinkage estimates of the probability that each subject broadly brackets, sorted from lowest to highest posterior mean. Error bars show a 95% Bayesian credible region. Note that for the Roommate's Dilemma on the top panel, almost all credible regions allow us to reject that the subject broadly brackets with probability greater than 50%. Similarly on the bottom panel almost all subjects are most likely to bracket narrowly in the Lottery Task.

Before turning to a richer model, we turn to the distribution of r^{RD}, r^{LT}, and γ. Moments of the transformed variables in Table 1.6 show that the average subject is risk-averse in both tasks (i.e. $r < 1$). This is relevant because the forgone utility due to bracketing narrowly increases as utility becomes more concave or convex (i.e. $|r-1|$ large). Considerable heterogeneity is in these variables is demonstrated through

Table 1.6.: Posterior means (standard deviations) of structural parameters.

Variable	r^G	r^L	γ	κ^G	κ^L
Transform	log	log	normal cdf	probit	probit
MEAN COEFFICIENTS β					
Constant	-0.9074	-0.4827	1.1945	-0.9838	-5.6477
	(0.2033)***	(0.0497)***	(0.5751)*	(0.2439)***	(4.1195)***
MOMENTS OF TRANSFORMED VARIABLES					
Mean	0.5708	0.6730	0.7624	0.1696	0.0485
	(0.0776)^a	(0.0306)^a	(0.1317)^a	(0.0562)^a	(0.0746)^a
Variance	0.3312	0.0851	0.0741	0.1376	0.0406
	(0.2124)^a	(0.0195)^a	(0.0602)^a	(0.0370)^a	(0.0604)^a
COVARIANCE MATRIX Σ					
r^G	0.6749		κ^G	1.0000^b	
	(0.2386)^a				
r^L	0.1186	0.1714	κ^L	-0.0948	1.0000^b
	(0.0494)**	(0.0334)^a		(0.5806)	
γ	0.4843	0.1130	1.5500		
	(0.4448)	(0.1147)	(1.3067)^a		
LOGIT CHOICE PRECISION					
λ^G	0.0445	λ^L	0.4621		
	(0.0224)^a		(0.0787)^a		

* $p < 0.050$, ** $p < 0.010$, *** $p < 0.001$,

[a] Variable restricted to be positive. Significance stars supressed

[b] Variable restricted to one. Significance stars and standard deviation suppressed

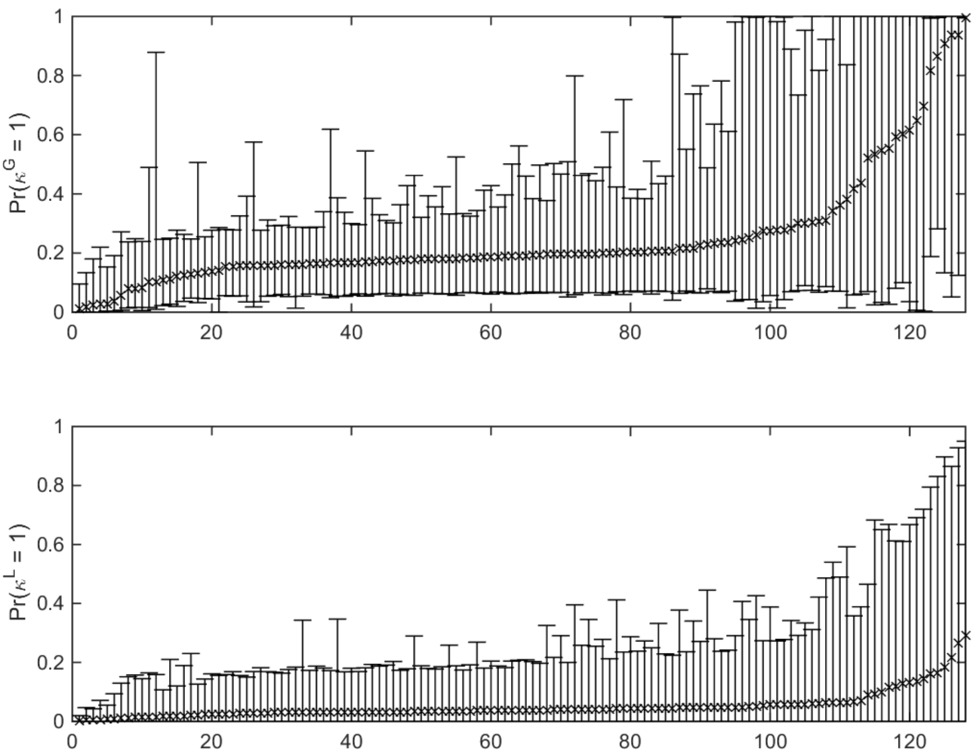


Fig. 1.5. Shrinkage estimates of the probability that each subject broadly brackets in the Roommate's Dilemma (top panel) and the Lottery Task (bottom panel), sorted from lowest to highest posterior mean. Crosses show posterior means, error bars show a 95% Bayesian credible region for these probabilities.

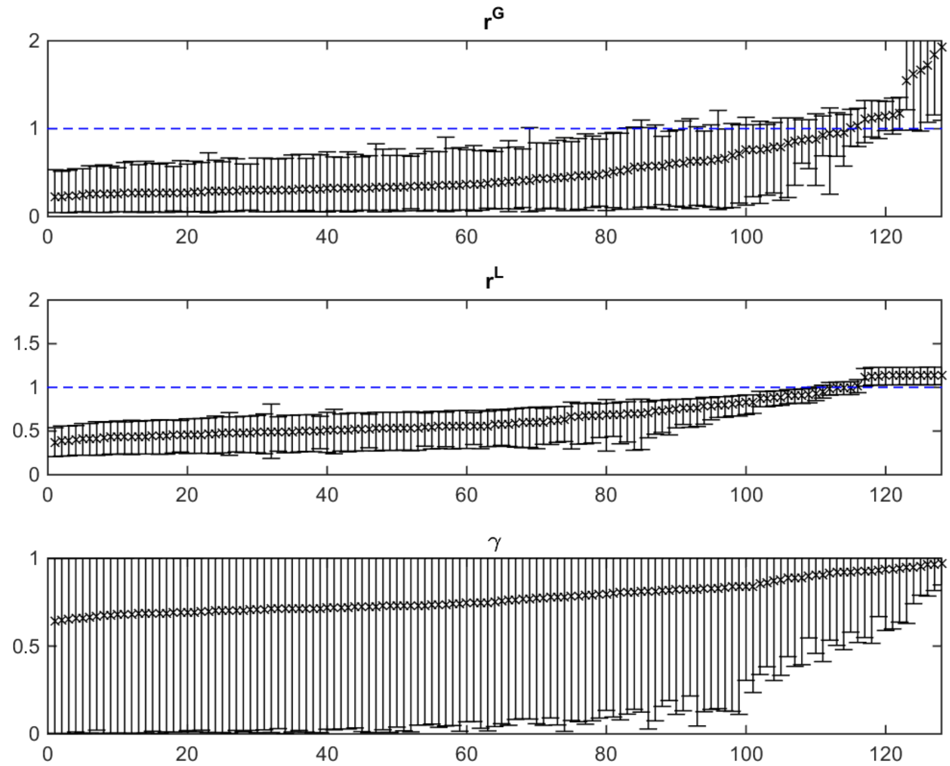


Fig. 1.6. Shrinkage estimates of behavioral parameters r^{RD}, r^{LT}, and γ, sorted from lowest to highest posterior mean. Crosses show posterior means, error bars show a 95% Bayesian credible region for these probabilities.

the estimated variance terms, and the shrinkage estimates of these variables shown in Figure 1.6. A correlation of 0.3217 (0.1147) is estimated between (logged) risk aversion in both tasks. We fail to reject a hypothesis that this correlation is positive (posterior probability = 0.9957).

Result 4 *In the Roommate's Dilemma, subjects who bracket narrowly stood to lose on average about 8% of their expected utility in certainty equivalent terms, compared to bracketing broadly. Since most subjects actually bracketed narrowly, subjects on average gave up approximately 7% of their expected utility due to narrow bracketing.*

Support: Figure 1.7 shows estimated losses that *could* have occurred due to bracketing errors (panel a), and estimated losses that occurred (panel b). The vertical axes in both panels show losses as a fraction of the optimized (broad bracketing) certainty equivalent, that is:

$$\frac{\text{CE assuming broad bracketing} - \text{CE assuming narrow bracketing}}{\text{CE assuming broad bracketing}} \tag{1.10}$$

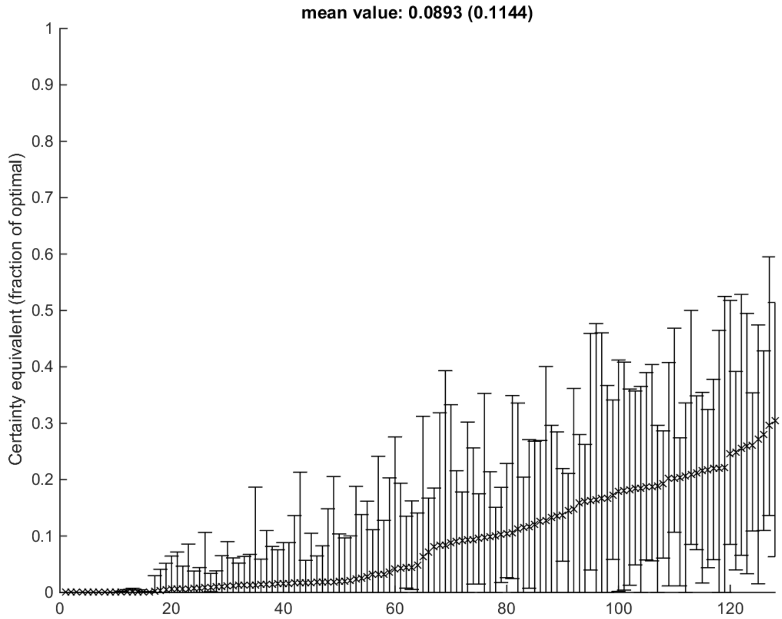
Averaging over all subjects, the potential losses due to bracketing errors (holding their estimated beliefs fixed) was about 8% of the certainty equivalent of acting optimally (see panel a). However this value hides considerable heterogeneity on the individual level. Subjects with risk neutral (or close to risk neutral preferences) stood to lose very little, and those with the most curvature in their utility functions [13] could lose about 30% of the certainty equivalent. Panel b shows the *actual* losses due to bracketing errors, which were about 7% of the certainty equivalent on average, but on the individual level were as high as 25%.

1.4 Conclusion

Economic experiments often produce results not well explained by standard economic models. Such deviations in individual settings are in their own right interesting,

[13] loosely speaking, either very risk averse or very risk loving subjects

(a) Potential



(a) Actual

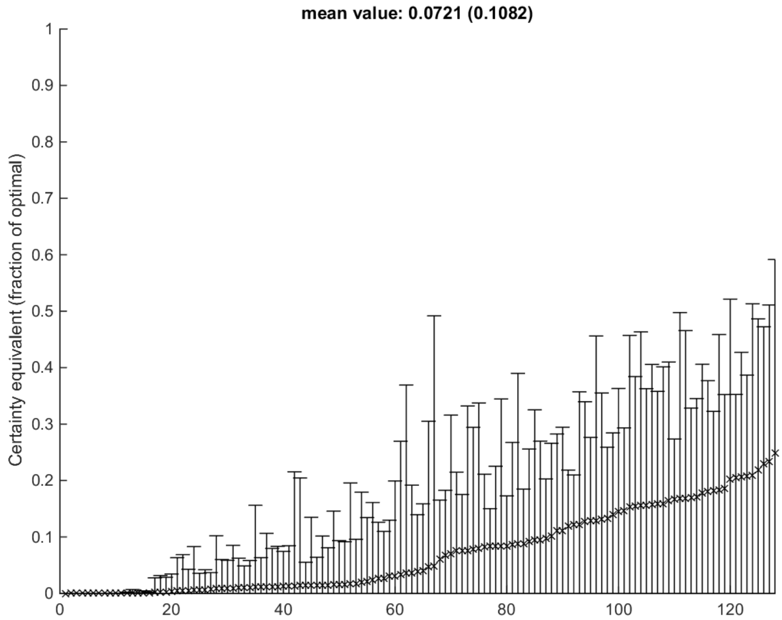


Fig. 1.7. Shrinkage estimates of forgone expected utility due to bracketing narrowly. Panel (a) shows potential losses that would have occurred if the subject had bracketed narrowly. Panel (b) shows the actual estimated losses. That is, in panel (b) losses were zero if the subject was (estimated to) broadly bracket.

but their relevance in games is not guaranteed until they are directly tested. I study one such observed deviation, choice bracketing, in a game. While this study, in line with Tversky and Kahneman (1981) and Rabin and Weizsacker (2009), finds that subjects' decisions in the absence of strategic considerations are better explained by narrow rather than broad bracketing, narrow and broad bracketing perform equally poorly in explaining aggregate treatment effects in the Roommate's Dilemma: treatments that should appear identical to a broad bracketer are played differently, and "games" that should appear identical are also played differently.

Structural modeling on the individual level finds that elicited risk preferences have some explanatory power for play in the Roommate's Dilemma, and the data seem to be best explained by a majority of narrowly bracketing subjects. This suggests that while assuming that agents best respond to *something* in games may not be too restrictive, assuming that they best respond to everything at the same time is probably wishful thinking for all but a small subset of the population.

While on average the estimated forgone utility of bracketing narrowly was small, some individuals stood to, and in fact did, forgo a significant fraction of their utility by making this error. How individuals bracket their decisions when acting strategically may therefore be an important consideration, both for predicting behavior, as well as for improving welfare.

2. MIXTURE MODELS OF BEHAVIOR AND NUISANCE PARAMETERS: A SEMI-PARAMETRIC BAYESIAN APPROACH

with Justin Tobias

2.1 Introduction

One of the most popular motivations for running economic experiments is that there is more than one model about how individuals behave in a particular environment.While in the past researchers may have considered these models as *competing*, aiming to identify the one that best explains behavior, research now acknowledges that each model could be "true" for different subsets of subjects. That is, each subject's behavior may be best described by one of the models being studied, but this model varies by subject. This motivates the mixture model, which assumes that the data are generated by more than one model, and the econometric task of estimating the fraction of subjects that behave according to each model.

One complication of this task is that models in economics are typically specified up to their parameters: the researcher wishes to distinguish between "Model A" and "Model B", not "Model A with given parameters θ^A" and "Model B with given parameters θ^B". As in Harrison and Rutström (2009) and Conte et al. (2011), for example, we may be interested in knowing the fraction of subjects that behave according to Expected Utility Theory and Prospect Theory, but may not care about an Expected Utility subject's degree of risk aversion. In the process of estimating these fractions, however, we must also estimate risk aversion parameters for all of our subjects (among others). Our assumption about how these individual-level parameters enter our econometric model is not inconsequential. Wilcox (2006), for example, demonstrates that failing to account for heterogeneity in learning models can produce biased estimates of the model parameters that favor one model of learning over another.[1]

Alternative models being studied in experiments frequently have different welfare implications. For example, in discussing expected utility and prospect theory, Thaler notes that:

> "If I had an important decision to make – whether to refinance my mortgage or invest in a new business – I would aim to make the decision in accordance with expected utility theory, just as I would use the Pythagorean theorem to estimate [the height of a] triangle. Expected utility is the right way to make decisions.
>
> With prospect theory, Kahneman and Tversky set out to offer an alternative to expected utility theory that had no pretense of being a useful guide to rational choice; instead, it would be a good prediction of the actual choices real people make. It is a theory about the behavior of Humans." Thaler (2015, p29)

We may therefore be much more interested in the model that an individual uses to make decisions, rather than the parameters that enter each model for each individual: in Thaler's context, individuals who use prospect theory to make decisions can be made better off if they can be helped to use the other model. Individuals who already use expected utility cannot be made better off (at least in this dimension), but their decisions could look very different depending on their risk aversion.

One complication with estimating mixing probabilities is that for some values of nuisance parameters, two or more of the assumed types could make similar decisions. If one is overly restrictive about the way the individual-level parameters enter the econometric model, it could be that the model achieves a better fit to the data by

[1] Wilcox (2006) demonstrates this with the task of identifying the "one best theory", not for mixture models. We use this example for exposition only and make no specific claims about their relevance in mixture models.

falsely classifying subjects into a different type. We show that mixture models that are incorrectly specified in this way can lead the researcher to falsely conclude that the data are generated by more than one type, when in fact all subjects are using the same decision rule.

We treat these individual-level parameters as nuisance parameters: variables which one must account for in one's estimation, but about which one does not directly care. Aiming to remain as agnostic about these as possible, we model them as draws from a finite mixture of (possibly multivariate) normal distributions. We then specify a data-generating process for an arbitrary experiment, and outline a Bayesian estimator for simulating the posterior distribution of its parameters. Of main interest are the parameters governing the mixing probabilities, which we model as a multinomial probit. This allows for statements to be made about how these mixing probabilities vary with observable characteristics of subjects, such as post-experiment survey responses, and treatment variables. Through the data augmentation process of the estimator, we also recover shrinkage estimates of the nuisance parameters and posterior probabilities that each subject behaves according to each model. This estimator is an improvement on the existing mixture model estimators used with experiments. Like Conte et al. (2011), we account for unobservable parameter heterogeneity at the individual level with a random coefficients specification, but with a more flexible distributional assumption placed on these parameters. As in Harrison and Rutström (2009), our estimator can identify how mixing probabilities change with observable subject characteristics, but does not assume that the nuisance parameters are a deterministic function of these same characteristics.

We then demonstrate the use of this estimator on two existing experimental datasets. We re-evaluate Harrison et al. (2010)'s assessment of the relationship between discount rates and smoking. We find that relaxing the econometric assumption about their nuisance parameters (risk aversion and discount rate) leads to much wider confidence intervals on the mixing probabilities: in our Bayesian framework the posterior distribution is close to the prior. We can conclude only that the population is at least 11% hyperbolic. Since the 95% credible interval for this probability is [2.2%, 45.4%] for non-smokers, and wider for smokers, we cannot say much about discounting behavior. Harrison et al. (2010), on the other hand, estimate a much tighter confidence interval of [18.3%, 36.0%].

Another advantage of our estimator is that it can accommodate more than two types. We demonstrate this by revisiting Andreoni and Vesterlund (2001)'s experiment, which documents differences in altruistic giving between male and female subjects. We augment their structural analysis by allowing subjects to place different weights on their partner's payoffs in three other-regarding utility functions: selfish, perfect substitutes, and perfect complements. We construct an overall demand curve for giving for males and females, and use it to assign a posterior probability to their claim that women are more generous when altruism is expensive, but men are more responsive to price. We reject their econometric assumption that non-selfish subjects place equal weight on their own and their partners' payoffs, and find economically significant heterogeneity in these weight parameters across subjects. This restriction, however, does not significantly change the type into which we classify most subjects.

2.2 Literature review

That there could be more than one relevant model of decision-making should not come as a surprise to experimental and behavioral economists. Alternative models are frequently discussed in studies of other-regarding preferences (Engelmann and Strobel, 2004; Bland and Nikiforakis, 2015), evaluating risky gambles (Rabin and Weizsacker, 2009), and discounting (Andersen et al., 2014; Coller et al., 2012), to name but a few.

Early treatments of alternative models of decision-making attempted to classify individual subjects into types. For example, in a modified dictator game experiment Andreoni and Miller (2002) classify subjects into three models of preferences over their own and another's monetary payoff: selfish, perfect substitutes, and perfect complements.Andreoni and Vesterlund (2001) extends this type of analysis by dividing subjects based on observable characteristics (sex), finding that men are more likely to have perfect substitutes preferences than women. An implicit assumption made in the analysis in these papers is that the non-selfish types place equal weight on their own and the other's payoff. They therefore subsume the problem of nuisance parameters in their analysis. Hey and Orme (1994) classified their subjects according to 11 different models of risky decision-making, estimating nuisance parameters separately for each subject. This process significantly reduces computational burden relative to many of the newer approaches discussed below, but does not capitalize on the information contained in the other subjects' decisions.

Harrison and Rutström (2009) demonstrate the implementation and value of mixture models in experiments by estimating the fraction of subjects who behave according to expected utility and prospect theory in a lottery choice experiment. They outline the mapping between individual likelihoods of each model, to the grand likelihood generated through the mixing probability. In our framework, there are five nuisance parameters: one in the expected utility model and four in the prospect theory model. Appealing to the existence of correlation between these nuisance parameters and subject characteristics, the authors model the nuisance parameters as a linear combination of subject characteristics X_i, that is:

$$\theta_i^\tau = X_i \delta^\tau \tag{2.1}$$

where θ_i^τ is subject i's nuisance parameter conditional on being type τ, X_i are their characteristics (e.g. age, sex, race, etc.), and δ^τ are the parameters being estimated. While this "linear regression" approach has appeal in that there is a very clear interpretation of δ^τ, in that the coefficient on (say) age is the marginal effect of age on θ_i^τ, what it masks is the strong assumption made about the relationship between characteristics and nuisance parameters: subjects with the same characteristics must have *exactly* the same nuisance parameters in *all* models. For example, if the only characteristic used were sex, then the econometrician assumes that all female subjects have the same level of risk aversion, probability of being an expected utility maximizer, probability weighting function, etc. Indeed, Harrison and Rutström (2009, footnote 27) acknowledge this and suggest a random effects specification, which would conform in our notation to:

$$\theta_i^\tau = X_i\delta^\tau + \epsilon_i^\tau, \tag{2.2}$$

with some specified distribution of ϵ_i^τ. Such a specification, albeit not in a mixture model, appears in von Gaudecker et al. (2011, see their Table 5), who also investigate risky decision making. These authors find that the mean ($X_i\delta^\tau$) and error (ϵ_i^τ) terms are *both* important in accounting for heterogeneity in choices across subjects. Harrison and Rutström (2009) admit that including both the deterministic and random components in the nuisance parameters of a mixture model would add "considerable complexity" to their estimator. For our paper we essentially add this random effects term for a very flexible distribution of ϵ_i^τ, but shut down the $X_i\delta^\tau$ component. Adding this back in within the Bayesian framework is not particularly difficult.

Again investigating risky decision-making, Conte et al. (2011) focus on a mixture model with expected utility and prospect theory decision-makers, but model heterogeneity of type (decision rule) and nuisance parameters as unobservable heterogeneity. This specification does not allow for parameters to vary by subjects' characteristics, but adds flexibility differently by modeling the nuisance parameters as draws from a multivariate normal distribution. Our procedure expands on Conte et al. (2011) by relaxing the multivariate normal assumption of the nuisance parameters, and allowing for the mixing probabilities to vary by subject characteristics (as in Harrison and Rutström, 2009).

2.3 Spurious type heterogeneity

At this point in the discussion, we consider the assumptions in Harrison and Rutström (2009) to be economically implausible: it is likely that subjects' preferences are not a deterministic function of a few demographic variables collected in an economic experiment. However we have not established that this assumption can lead to false conclusions about the prevalence of types in the population. While we consider the estimator used in Conte et al. (2011) to be economically more plausible with respect to the nuisance parameters, this technique does not allow for the mixing probability to vary by observables, and so cannot be used to identify subsets of the population who are significantly more "behavioral" than others, outside of simple partitionings of the data. In this section we demonstrate that mixture models that assume a deterministic relationship between nuisance parameters and observables are very sensitive to the number of types included, which in this setup is equivalent to adding some simple nuisance parameter heterogeneity. We show using simulation results that this technique can lead the researcher to conclude that there is a significant fraction of both types in the sample, even when the data are generated with all subjects using the same decision rule.

2.3.1 Harrison and Rutström (2009)

Mixture models estimated with the Harrison and Rutström (2009) method assume that subjects' nuisance parameters and mixing probabilities depend on their observable characteristics only. One such application of this method is Harrison et al. (2010), in which the authors use an experiment to elicit risk and time preferences, and investigate the relationship between these elicited preferences, the type of discounting a subject uses, and whether the subject is a smoker. Of particular interest to policymakers could be the fraction of smokers compared to non-smokers that are hyperbolic discounters. If smokers are disproportionately hyperbolic discounters, then one could attribute the decision to smoke as an irrational one, rather than the result of rational choice with different time and risk preferences. We therefore investigate the sensitivity of the mixing parameters to alternative specifications within the Harrison and Rutström (2009) framework. Table 2.1 shows the results these alternatives.

Column 1 replicates Table 5 of Harrison et al. (2010). Here, they assume that there are two types: exponential and hyperbolic discounters. The nuisance parameters in this model are CRRA risk aversion r, and discount rate δ. They assume that risk aversion does not depend on type, but discount rate depends on type. As their only explanatory variable in this estimation is `smoke`, a categorical variable equal to one if the subject is a smoker, and zero otherwise, we can represent their assumption about r and δ as:

$$r_i = \beta_0^r + \beta_1^r \text{smoke}_i \tag{2.3}$$

$$\delta_i = \beta_0^\delta + \beta_1^\delta \text{smoke}_i + \beta_3^\delta I(\tau_i = 1) + \beta_4^\delta \text{smoke}_i I(\tau_i = 2) \tag{2.4}$$

As in their Table 5, we transform the actual parameters estimated into risk aversion by smoking status, and discount rate by both smoking status and type. The conclusions from their estimation are that there are significant fractions of both exponential and hyperbolic types in the population (they separately reject hypotheses that the mixing probabilities are equal to zero and one. $p < 0.001$ in both cases.), and that smokers are no more likely to make hyperbolic decisions than are non-smokers ($p = 0.349$).

We begin our extension by showing that alternative assumptions about the types can generate better fits of the data. These alternative specifications are shown in columns 2 and 3. In column 2, we assume that both types are exponential. The purpose of this specification is to demonstrate the potential of drawing a false conclusion about types by assuming the specification in column 1. Column 2 estimates that there is a significant fraction of "Type 1"s and "Type 2"s in the population, however this specification invites a different conclusion: as the assumed model for decision-making was the same across these types (i.e. both types are exponential), one could use this as evidence for heterogeneity in *preferences*. That is, between-subject variation in the

Table 2.1.: Mixture model estimation from Harrison et al. (2010) experiment on discounting behavior and smoking (column 1), and alternative specifications.

	(1)	(2)	(3)	(4)	(5)
Risk aversion – r					
non-smoker	0.792***	0.765***	0.797***	0.781***	0.768***
	(0.0527)	(0.0509)	(0.0524)	(0.0501)	(0.0494)
- smoker	0.837***	0.814***	0.842***	0.845***	0.832***
	(0.0534)	(0.0548)	(0.0525)	(0.0503)	(0.0528)
Discounting – δ					
Type 1					
non-smoker	0.0590***	0.0570***	0.0590***	0.0359***	0.0333***
	(0.00813)	(0.00743)	(0.00820)	(0.00558)	(0.00487)
smoker	0.0730***	0.0718***	0.0758***	0.0503*	0.0391***
	(0.00971)	(0.00948)	(0.0103)	(0.0243)	(0.00992)
Type 2					
non-smoker	0.327***	0.308***	0.324***	0.211***	0.202***
	(0.0439)	(0.0340)	(0.0429)	(0.0203)	(0.0184)
smoker	0.267***	0.260***	0.269***	0.187***	0.168***
	(0.0492)	(0.0405)	(0.0478)	(0.0358)	(0.0266)
Type 3					
non-smoker				0.528***	0.448***
				(0.0817)	(0.0556)
smoker				0.355*	0.350***
				(0.139)	(0.0628)
Mixing probabilities – π					
Type 2					
non-smoker	0.272***	0.321***	0.275***	0.456***	0.460***
	(0.0450)	(0.0437)	(0.0442)	(0.0407)	(0.0373)
smoker	0.379***	0.416***	0.354***	0.459*	0.503***
	(0.106)	(0.0924)	(0.0983)	(0.200)	(0.0857)
Type 3					
non-smoker				0.107***	0.144***
				(0.0184)	(0.0206)
smoker				0.107*	0.157***
				(0.0522)	(0.0379)
Choice precision – μ					
Risk aversion	0.0673***	0.0749***	0.0658***	0.0687***	0.0725***
	(0.0161)	(0.0156)	(0.0161)	(0.0149)	(0.0149)
Discounting	0.00956**	0.0104***	0.00920**	0.00675**	0.00695**
	(0.00332)	(0.00307)	(0.00323)	(0.00225)	(0.00211)
N	23008	23008	23008	23008	23008
ll	-12241.9	-12198.8	-12245.8	-12159.4	-12132.4
Type1	Exponential	Exponential	Hyperbolic	Exponential	Exponential
Type2	Hyperbolic	Exponential	Hyperbolic	Exponential	Exponential
Type3	(absent)	(absent)	(absent)	Hyperbolic	Exponential

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

data is largely driven by subjects having different rs and δs, not by different decision rules. Even more alarmingly, we note that column 2 achieves a greater likelihood than column 1 (-12198.8 vs. -12241.9). It appears that going from one exponential type to two exponential types buys more explanatory power than adding a hyperbolic type. In the interest of including the exhaustive set of Harrison and Rutström (2009) style estimations assuming two types, we also include column 3, which assumes that there are two hyperbolic types only. By the log-likelihood metric, this specification does worse than the one reported in Harrison et al. (2010).

We now turn to demonstrating that this type of estimation may lead to false conclusions about the mixing probabilities. In column 4, we assume that there are three types: two exponential and one hyperbolic. We are interested in comparing the estimated fraction of hyperbolic types here to the corresponding number in column 1. In column 4 we estimate that a much smaller fraction of decisions, about 11% vs. 30% in column 1, are hyperbolic. Depending on one's opinion of what constitutes an economically significant fraction of irrational types, one's conclusion from these models could therefore be very different. As this model nests column 1, we can use the likelihood ratio test to reject the 2-type assumption in favor of adding another exponential type.[2] For the sake of completeness, we also include column 5, which assumes three exponential types, and no hyperbolic types. Comparing the log likelihood to column 4, it appears that swapping the hyperbolic type for an exponential type buys additional explanatory power.

It is important to note here that we do not wish the reader to conclude from our analysis in Table 2.1 that column 1, and hence Harrison et al. (2010), overstates the prevalence of hyperbolic discounting in the population. Rather, we view our analysis as demonstrating how the estimate of the mixing probabilities is very susceptible to the assumptions made about the nuisance parameters: *all* of the mixing probabilities in Table 2.1 should be taken with a grain of salt. If researchers are interested in

[2] $\chi^2_2 = 2(12241.9 - 12159.4) = 165.0$, $p < 0.001$.

making statements about mixing probabilities, we therefore encourage them to think carefully about how nuisance parameters enter into their econometric specification.

2.3.2 A pathological example

In the previous section, we demonstrated that alternative treatments of the nuisance parameters can have large effects on conclusions drawn from the estimation results. One plausible explanation of these varying results is that there is significant unobservable heterogeneity in the nuisance parameters. Instead of the assumption made in (2.3) and (2.4), consider the alternative assumption that:

$$\theta_i^\tau = \mu^\tau + \epsilon_i^\tau \tag{2.5}$$

where μ^τ is the expected value of the nuisance parameter, and ϵ_i^τ is an iid, mean zero error term. If ϵ_i^τ is drawn from a degenerate distribution, then the Harrison and Rutström (2009) estimator is correctly specified, or at least the nuisance parameters enter the model correctly. On the other hand, if ϵ_i^τ is not drawn from a degenerate distribution, then this estimator makes a simplifying assumption.

This example demonstrates a possible cause of the varying results in Table 2.1. Consider two competing theories about how subjects make decisions y_i based on treatment variable t. In particular, we are interested in two potentially relevant theories:

1. Type A "linear" subjects maximize their utility by choosing $y^* = \theta_i^A t$
2. Type B "log" subjects maximize their utility by choosing $y^* = \theta_i^B \log(1+t)$

Since our subjects are humans, they make mistakes in their optimization, which we model as their actual choices being their optimal choices plus a normal error term:

$$y_{i,t} = y_{i,t}^* + \epsilon_{i,t}^y, \quad \epsilon_{i,t}^y \text{ iid} \sim N(0, \sigma^2) \tag{2.6}$$

Each subject makes T decisions for $t = t_1, t_2, \ldots, t_T$.

We wish to estimate π^A, the fraction of linear subjects in our population. The challenge in separating out these two types is that for $t \approx 1$, $\log(1+t) \approx t$. It is therefore difficult to distinguish between them if the experiment only studies treatments for t near one. This, of course, *should* be addressed with better experimental design, in this case studying t sufficiently far away from 1. We focus on the case where the experiment is fixed, and so the task remains solely an econometric one.

In our pathological example, there are only linear subjects in the population, but half of these subjects have $\theta_i^A = 1$, and the other half have $\theta_i^A = 2$. In terms of (2.5), this corresponds to:

$$\mu^A = 1.5, \quad \epsilon_i^A = \begin{cases} -0.5 & \text{with probability } 0.5 \\ 0.5 & \text{with probability } 0.5 \end{cases} \tag{2.7}$$

If we let π^A be the fraction of linear types in our sample, we would like whatever technique we apply to our data to produce an estimate $\hat{\pi}^A$ close to 1. That is, we may be concerned that we falsely conclude that there is a significant fraction of log subjects in our sample.

Suppose that we apply an estimator that assumes no heterogeneity in the nuisance parameters, and that we are lucky enough to know σ^2, the decision error term. An important quantity is the (log) odds ratio of a subject's decisions conditional on being each type, which in this case is:

$$OR_i = -\sum_{k=1}^{T} \left(\frac{y_{i,k} - \theta^A t}{\sigma} \right)^2 + \sum_{k=1}^{T} \left(\frac{y_{i,k} - \theta^B \log(1+t)}{\sigma} \right)^2 \tag{2.8}$$

As this number increases, it is more likely that subject i is the linear type. Taking a Taylor expansion of the log term yields:

$$\sigma^2 OR_i = -\sum_{k=1}^{T} \left(y_{i,k} - \theta^A t\right)^2 + \sum_{k=1}^{T} \left(y_{i,k} - \theta^B (t + o(t^2))\right)^2 \tag{2.9}$$

$$= -\sum_{k=1}^{T} \left(y_{i,k} - \theta^A t\right)^2 + \sum_{k=1}^{T} \left(y_{i,k} - \theta^B t - \theta^B o(t^2))\right)^2 \tag{2.10}$$

$$= -\sum_{k=1}^{T} \left[(y_{i,k} - \theta^A t)^2 - (y_{i,k} - \theta^B t)^2\right] \tag{2.11}$$

$$- \theta^B o(t^2) \sum_{k=1}^{T} \left[2(y_{i,k} - \theta^B t) - \theta^B o(t^2)\right]$$

where $o(t^2) = \log(1 + t) - t$ is the error associated with the linear approximation of $\log(1+t)$. In (2.11), for t close to 1, the first term dominates. In this case classification of subjects into types is driven almost exclusively by differences between $\theta^A t$ and $\theta^B t$. Therefore, with our distribution in (2.7), the best our estimator can do is fit half of the data to a linear function of t, and the other half to a log function. By assumption, we have not given our estimator the opportunity to fit the data to two linear functions.

2.3.3 Simulation results

We now further explore the implications of (2.11) with Monte Carlo simulation. We consider the implications making incorrect assumptions about the distribution of types and nuisance parameters, and show that the correctly specified models accurately estimate the mixing probability. We simulate data from the hypothetical experiment described above, based on two cases:

1. There is one "linear" type A whose optimal choice is $y^* = t$ (i.e. $\theta^A = 1$), and one "log" type B, whose optimal choice is $y^* = \theta^B \log(1 + t)$.

2. There are two linear types. Type A's optimal choice is $y^* = t$ (i.e. $\theta^A = 1$), and type B"s is $y^* = \theta^B t$

Table 2.2.: Monte Carlo simulation results from example described in Section 2.3. Values show the mean estimate of the fraction of linear types. Root mean squared errors are in parentheses. These numbers were constructed with a simulation size of 5,000, with each run simulating 100 subjects making 20 decisions each.

		True value of θ^B			
Actual B	Assumed B	0.5	0.7	1.3	2.0
log	log	0.5003	0.5016	0.4999	0.5000
		(0.0127)	(0.0509)	(0.0297)	(0.0000)
linear	log	0.5037	0.6119	0.8047	0.5001
		(0.4965)	(0.3970)	(0.2217)	(0.4999)
linear	linear	0.5005	0.5008	0.5021	0.5000
		(0.0122)	(0.0577)	(0.0569)	(0.0002)

We perform this simulation for $\theta^B \in \{0.5,\ 0.7,\ 1.3,\ 2.0\}$, simulating data from an experiment with 100 subjects, exactly half of which are type A. Each subject makes $T = 20$ decisions for t evenly spaced between -0.3 and 0.3. We use a decision error variance of $\sigma^2 = 0.01$, which we assume is known to the econometrician (in order to speed up the simulation). Table 2.2 shows the results of this simulation.

The first row of this table reports the results from a correctly specified model where there is one linear type and one log type. The estimator pins down the mixing probability well, on average correctly estimating that there are equal fractions of both types. RMSE increases as θ^B approaches 1, which is to be expected because $\log(1 + x) \approx x$ for $x \approx 0$, and hence the experiment does not classify subjects as cleanly.

The second row shows the result of simulating data with both types being linear, but assuming that type B is log. Here the "true" mixing probability is still 50%, however we note that *estimating* π^A close to 0.5 would result in the erroneous conclusion that there are close to equal proportions of log and linear types in the population. If a researcher were to blindly interpret $\hat{\pi}^A$ this way, one would hope that this value was not significantly different to one, and hence conclude that all subjects are linear. We have therefore reported RMSE for this row by calculating deviations from a true value of 1, rather than 0.5. In reality, however, the simulation reveals that the estimate of π^A frequently leads to the wrong conclusion.

The third row of Table 2.2 shows results for the type B linear DGP with the correct specification. As with the first row, as researchers we hope that the estimated mixing probability is close to 0.5, indicating in this case that there are roughly equal proportions of linear types, who can be differentiated by their nuisance parameter. This is indeed the case.

2.4 A more flexible data-generating process

In the previous section, we showed that incorrect assumptions about the nuisance parameters can propagate into incorrect conclusions about the mixing probabilities. One solution to this, as used in von Gaudecker et al. (2011), is to individually estimate the nuisance parameters for each subject. However to achieve accurate estimates of these, we often need long, expensive experiments. Instead, we use a hierarchical specification, where nuisance parameters are drawn from a specified population distribution. In Conte et al. (2011) this is a multivariate normal distribution. We relax this assumption further in assumed data-generating process described below.

In an experiment, subject i takes J_i actions $a_i = \{a_{i,j}\}_{j=1}^{J_i}$.

There are T *types* of subject. Each type is a functional form describing a subject's behavior. We model this as a likelihood function $l_\tau(a, \mid z, \theta_i^\tau, \gamma)$, where a is the action taken by a subject, z is a treatment condition, θ_i^τ is a set of behavioral parameters that describe subject i's behavior conditional on being type τ, and γ is a set of parameters common to all types.

We model the determination of types as a (iff $T \geq 3$, multinomial) probit, and let β represent this parameter of interest. Σ is the covariance matrix of the underlying latent process (discussed below) determining types. We impose the identifying restriction $\Sigma_{T-1,T-1} = 1$.

Conditional on being type τ, parameter θ_i^τ is an iid draw from distribution $f_\tau(\theta)$. As the aim of this study is to estimate β, and not θ, we wish to remain as agnostic as possible about this distribution. Hence, we opt for the semi-parametric approach of letting $f_\tau(\theta)$ be a mixture of C_τ (possibly multivariate) normal distributions. Specifically:

$$f_\tau(\theta) = \sum_{c=1}^{C_\tau} \psi_{\tau,c}\phi\left(\theta; \mu_{\tau,c}, \nu_{\tau,c}\right) \tag{2.12}$$

where $(\mu_{\tau,c}, \nu_{\tau,c})$ are the mean and variance for component c in type τ, and ψ_τ is the vector of mixing probabilities for type τ.

2.5 Bayesian implementation

We specify priors on all parameters $(\beta, \Sigma, \gamma, \psi, \mu, \nu)$ and estimate their posterior distribution using data augmentation and Gibbs sampling. This technique involves first defining latent variables for the censored data, then deriving (or otherwise simulating) the posterior distribution of all parameters, conditional on all others.

2.5.1 Latent variables

We adopt the latent variables approach to the multinomial probity for τ_i:

$$\tau_i^* \mid \beta, \gamma, \psi, \mu, \nu, \theta \sim \; iid\, \mathcal{N}(X\beta, \Sigma) \tag{2.13}$$

$$\tau_i = \begin{cases} 1 & \text{if } \arg\max_k\{\tau_{i,k}^*\} = 1 \\ \vdots & \quad\vdots \\ t & \text{if } \arg\max_k\{\tau_{i,k}^*\} = t \\ \vdots & \quad\vdots \\ T-1 & \text{if } \arg\max_k\{\tau_{i,k}^*\} = T-1 \\ T & \text{otherwise} \end{cases} \tag{2.14}$$

Furthermore, as τ_i is not observed, it is also a latent variable.

Additionally, we let $c_{i,\tau}$ indicate the component that subject i's parameters were drawn from.

2.5.2 Augmented likelihood function and posterior distribution

In addition to the model parameters $(\beta, \Sigma, \gamma, \psi, \mu, \nu)$, we augment the data with $\{\tau_i^*, \tau_i, c_{i,\tau_i}, \theta_i\}_{i=1}^N$. Therefore, we can define the augmented likelihood function of observing data $\{a_{i,j}\}$:

$$\begin{aligned} & p(a, \tau^*, \tau, c, \theta \mid \beta, \Sigma, \gamma, \psi, \mu, \nu) \\ & \quad \propto \prod_{i=1}^N I(\tau_i = \arg\max_k\{\tau_{i,k}^*\})\phi(\tau_i^*; X_i\beta, \Sigma) \\ & \qquad \times \psi_{\tau, c_{i,\tau_i}} \phi(\theta_i^\tau; \mu_{c_{i,\tau_i}}^\tau, \nu_{c_{i,\tau_i}}^\tau) \times l_{\tau_i}(a_i \mid z, \theta_i^{\tau_i}, \gamma) \end{aligned} \tag{2.15}$$

where $I(\cdot)$ is the indicator function. Combining this with priors on the relevant variables yields the the posterior distribution:

$$\begin{aligned} &p(\beta, \Sigma, \gamma, \psi, \mu, \nu, \tau^*, \tau, c, \theta \mid a) \\ &\quad \propto \prod_{i=1}^{N} I(\tau_i = \arg\max_k \{\tau_{i,k}^*\}) \phi(\tau_i^*; X_i\beta, \Sigma) \\ &\qquad \times \psi_{\tau, c_{i,\tau_i}} \phi(\theta_i^\tau; \mu_{c_{i,\tau_i}}^\tau, \nu_{c_{i,\tau_i}}^\tau) \times l_{\tau_i}(a_i \mid z, \theta_i^{\tau_i}, \gamma) \\ &\qquad \times p(\beta)p(\Sigma)p(\gamma)p(\psi)p(\mu, \nu) \end{aligned} \tag{2.16}$$

We adopt the following prior distributions, which admit conjugate posterior distributions.

Conditional posterior distribution of parameters determining types

Parameters (β, Σ) determine the probability that a subject behaves according to each type. Latent parameters (τ^*, τ) govern this process.

Latent type parameters (τ^*, τ): In order to draw these parameters conditional on all others, we use the following result:

$$\begin{aligned} &p(\tau_i^*, \tau_i \mid F) = p(\tau_i^* \mid \tau_i, F) p(\tau_i \mid F) \\ &\text{where:} \quad F = \beta, \Sigma, \gamma, \psi, \mu, \nu, \tau_{-i}^*, \tau_{-i}, c, \theta, a \end{aligned} \tag{2.17}$$

In the first step, we draw $\tau_i \mid F$:

$$p(\tau_i \mid F) \propto l_\tau(a_i \mid z, \theta_i^\tau, \gamma) p(\tau_i \mid F - a_i) \tag{2.18}$$

Here $p(\tau \mid F - a_i)$ is the probability subject i is type τ, but unconditional on their actions. In the simplest mixture model case where $T = 2$, this would be $p(\tau_i = 1 \mid$

$F - a_i) = \Phi(X_i\beta)$, the probit prediction probability, in the multivariate case this is the multivariate probit analog:

$$p(\tau_i = k \mid F - a_i) = \begin{cases} p(k = \arg\max_j\{\tau_j^*\} \cap \tau_j^* \geq 0) & \text{if } k \in \{1, 2, \ldots, T-1\} \\ p(\max_j\{\tau_j^*\} < 0) & \text{if } k = T \\ 0 & \text{otherwise} \end{cases} \tag{2.19}$$

As τ_i can only take on T possible values, we evaluate (2.18) at all of these, and divide by the sum to get the desired pmf of τ_i, then draw from this.

Given τ_i, τ_i^* is a draw from a truncated multivariate normal distribution with mean $X_i\beta$ and covariance Σ such that element τ_i is the maximum, unless $\tau_i = T$, in which case all elements are less than zero.

Probit parameters (β, Σ)**:** We use the normal-inverse Wishart prior on these parameters:

$$\beta \sim \mathcal{N}\left(\underline{\mu}_\beta, \underline{V}_\beta\right) \tag{2.20}$$

$$\Sigma^{-1} \sim \text{Wishart}\left([\underline{\rho}\underline{R}]^{-1}, \underline{\rho}\right) I(\Sigma_{T-1,T-1} = 1) \tag{2.21}$$

Inspection of (2.16) yields that the conditional posterior distribution of these parameters is proportional to:

$$p(\beta, \Sigma \mid \gamma, \psi, \mu, \nu, \tau^*, \tau, c, \theta, a) \propto \prod_{i=1}^{N} \phi(\tau_i^*; X_i\beta, \Sigma)p(\beta)p(\Sigma) \tag{2.22}$$

which yields the standard results:[3]

$$\beta \mid \Sigma^{-1}, \gamma, \psi, \mu, \nu, \tau^*, \tau, c, \theta, a \sim \mathcal{N}(D_\beta d_\beta, D_\beta) \tag{2.23}$$

$$\text{where:} \quad D_\beta = \left(\sum_{i=1}^{N} Z_i' \Sigma^{-1} Z_i + V_\beta^{-1} \right)^{-1}$$

$$d_\beta = \sum_i Z_i' \Sigma^{-1} \tau_i^* + \underline{V}_\beta^{-1} \underline{\mu}_\beta \tag{2.24}$$

and:

$$\Sigma^{-1} \mid \beta, \gamma, \psi, \mu, \nu, \tau^*, \tau, c, \theta, a \tag{2.25}$$

$$\sim \text{Wishart} \left(\left[\underline{\rho} \underline{R} + \sum_i^N (\tau_i^* - Z_i \beta)(\tau_i^* - Z_i \beta)' \right]^{-1}, N + \underline{\rho} \right) \tag{2.26}$$

Parameters governing the distributions of θ_i^τ

Component parameters c_i^τ: The "priors" for these variables are captured by the component mixing parameters ψ:

$$p(c_i^\tau \mid \beta, \Sigma^{-1}, \gamma, \psi, \mu, \nu, \tau^*, \tau, c_{-c_i^\tau}, \theta) \propto \psi_c^\tau \phi(\theta_i^\tau; \mu_c^\tau, \nu_c^\tau) l_\tau(a_i \mid z, \theta_i^\tau, \gamma) \tag{2.27}$$

By noting that c_i^τ can only take on integers between 1 and C_τ, this conditional distribution can be calculated by evaluating the right-hand side of (2.27) at these integers, and dividing by the sum of the results.

[3]See for example Koop et al. (2007) exercise 14.7 for the derivation. Here we make the transform:

$$Z_i = \begin{bmatrix} X_i & 0 & \dots & 0 \\ 0 & X_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & X_i \end{bmatrix}$$

Component mixing parameters ψ: By using the prior:

$$\psi_\tau \sim \text{Dirichlet}\left(\underline{\alpha}_{\tau,1}, \underline{\alpha}_{\tau,2}, \ldots, \underline{\alpha}_{\tau,C_\tau}\right) \tag{2.28}$$

the posterior distribution is proportional to:

$$p(\psi_\tau \mid \beta, \Sigma^{-1}, \gamma, \psi_{-\tau}, \mu, \nu, \tau^*, \tau, c, \theta, a) \propto p(\psi_\tau) \prod_{\forall i:\tau_i=\tau} \psi_{\tau, c_{i,\tau_i}} \tag{2.29}$$

it follows that:

$$\psi_\tau \mid \beta, \Sigma^{-1}, \gamma, \psi_{-\tau}, \mu, \nu, \tau^*, \tau, c, \theta, a \sim \text{Dirichlet}\left(\overline{\alpha}_{\tau,1}, \overline{\alpha}_{\tau,2}, \ldots, \overline{\alpha}_{\tau,C_\tau}\right) \tag{2.30}$$

where: $\overline{\alpha}_{\tau,k} = \sum_{i=1}^{N} I(\tau_i = \tau) I(c_i = k)$

Component distribution parameters (μ, ν): The section of (2.16) proportional to each components' parameters (μ_c, ν_c) is:

$$p(\mu_c^\tau, \nu_c^\tau \mid \beta, \Sigma^{-1}, \gamma, \psi, \mu_{-c}, \nu_{-c}, \tau^*, \tau, c, \theta, a) \propto p(\mu_c^\tau, \nu_c^\tau) \prod_{i:c_i=c, \tau_i=\tau} \phi(\theta_i^{\tau_i}, \mu_c^\tau, \nu_c^\tau) \tag{2.31}$$

If we use a normal-inverse Wishart prior for (μ_c^τ, ν_c^τ):

$$\mu_c^\tau \sim \mathcal{N}\left(\underline{M}_\mu, \underline{V}_\mu\right) \tag{2.32}$$

$$(\nu_c^\tau)^{-1} \sim \text{Wishart}\left((\underline{A}_c^\tau)^{-1}, \underline{b}_c^\tau\right) \tag{2.33}$$

then the posterior distributions are:

$$\mu_c^\tau \mid \beta, \Sigma^{-1}, \gamma, \psi, \mu_{-c}, \nu, \tau^*, \tau, c, \theta, a \sim \mathcal{N}(D_\mu d_\mu, D_\mu) \tag{2.34}$$

$$\text{where: } D_\mu = \left[\underline{V}_\mu^{-1} + (\nu_c^\tau)^{-1} \sum_i I(\tau_i = \tau) I(c_i^\tau = c)\right]^{-1} \tag{2.35}$$

$$d_\mu = \underline{V}_\mu^{-1} \underline{M}_\mu + (\nu_c^\tau)^{-1} \frac{\sum_i I(\tau_i = \tau) I(c_i^\tau = c) \theta_i^\tau}{\sum_i I(\tau_i = \tau) I(c_i^\tau = c)} \tag{2.36}$$

and:

$$(\nu_c^\tau)^{-1}\,|\beta,\Sigma^{-1},\gamma,\psi,\mu,\nu_{-c},\tau^*,\tau,c,\theta,a \sim \text{Wishart}\left((\overline{A}_c^\tau)^{-1},\overline{b}_c^\tau\right) \tag{2.37}$$

where:

$$\overline{A}_c^\tau = \underline{A}_c^\tau + \sum_i I(\tau_i=\tau)I(c_i^\tau=c)(\theta_i^\tau-\mu_c^\tau)(\theta_i^\tau-\mu_c^\tau)' \tag{2.38}$$

$$\overline{b}_c^\tau = \underline{b}_c^\tau + \sum_i I(\tau_i=\tau)I(c_i^\tau=c) \tag{2.39}$$

Parameters without a known conditional posterior:

Without being more specific about how common parameter γ and individual behavioral parameters θ_i^τ enter into the decision-making problem, conditional posterior distributions cannot be derived. In these cases we opt for rejection sampling using the Metropolis-Hastings algorithm, which requires only that the target density be known up to a constant of proportionality. Inspection of (2.16) yields that:

$$p(\theta_i^\tau \mid \beta,\Sigma^{-1},\gamma,\psi,\mu,\nu,\tau^*,\tau,c,\theta_{-i},a) \propto \phi(\theta_i^\tau;\mu_{c_i}^\tau,\nu_{c_i}^\tau)l_\tau(a_i \mid z,\theta_i^\tau,\gamma) \tag{2.40}$$

$$p(\gamma \mid \beta,\Sigma^{-1},\gamma,\psi,\mu,\nu,\tau^*,\tau,c,\theta,a) \propto p(\gamma)\prod_{i=1}^N l_{\tau_i}(a_i \mid z,\theta_i^{\tau_i},\gamma) \tag{2.41}$$

It should be noted, however, that in some cases these Metropolis-Hastings steps may be replaced by more direct methods with the use of an intelligently chosen prior and/or data augmentation process. Such an example can be found in Section 2.6.1

2.6 Applications

In this section we apply our estimator to existing datasets. First, we use Andreoni and Vesterlund (2001) to demonstrate the simple case where there is no more than

Table 2.3.: Budgets used in the Andreoni and Vesterlund (2001) experiment (their Table I).

Budget	Token endowment	Hold value	Pass value
1	40	1	3
2	60	1	2
3	75	1	2
4	60	1	1
5	100	1	1
6	60	2	1
7	75	2	1
8	40	3	1

one nuisance parameter per type. Secondly, we apply our estimator to Harrison et al. (2010).

2.6.1 Andreoni and Vesterlund (2001)

Andreoni and Vesterlund (2001) investigate whether men and women behave differently in a modified dictator game. In this experiment, subjects were given a budget of tokens, and decided how to allocate these between themselves and another person who could not affect the outcome. Using a within-subjects design, subjects were presented with a series of these decisions, which varied in both the number of tokens, and the value of tokens in points (worth US$0.10 each) to themselves and the other person. In total subjects made eight decisions, with the token values and endowments summarized in Table 2.3. For example in "Budget 1", the decision-maker was asked to "Divide 40 tokens: Hold ___ at 1 point each, and Pass ___ at 2 points each". If the decision-maker held all 40 tokens, she would earn 40 points and the other would earn nothing. If she held half of the tokens, she would earn 20 points and the other would earn 60 points. All subjects made decisions for all budgets. Pairs and roles (decision-maker and receiver) were randomly assigned at the end and one randomly selected budget was paid.

Their main result is that "when altruism is expensive, women are kinder, but when it is cheap, men are more altruistic" (Andreoni and Vesterlund, 2001, abstract), however we focus on their classification of subjects into types, as reported in their Table III. The task is essentially the consumers' problem: the decision maker (henceforth "the self") must decide how to allocate their endowment of tokens between themself and another individual (henceforth "the other"). Subjects make eight such decisions with different endowments, and importantly, different "prices". That is, the value of tokens to the self and the other varies. Assuming homothetic preferences, which is a maintained assumption in Andreoni and Vesterlund (2001)'s structural analysis, the consumer's problem reduces to:

$$\max_{t\in[0,1]} u_i(t,(1-t)/p_k) \tag{2.42}$$

where t is the fraction of the endowment of the tokens kept, and p_k is price of increasing the other's income by one unit in treatment k.

In their Table III, subjects are classified using a minimum-distance estimator into three types: Selfish subjects maximize $U(\pi_s,\pi_o)=\pi_s$, Perfect Compliments (Leontief) subjects maximize $U(\pi_s,\pi_0)=\min\{\pi_s,\pi_0\}$, and Perfect Substitutes subjects maximize $U(\pi_s,\pi_o)=\pi_s+\pi_o$. In the interest of adding some meaningful heterogeneity within these types, we introduce nuisance parameters for the weight placed on the other's payoff: Selfish subjects, as before, maximize $U_i^S(\pi_s,\pi_o)=\pi_s$; Perfect Compliments (Leontief) subjects maximize $U^{PC}(\pi_s,\pi_0)=\min\{\theta_i^{PC}\pi_s,\pi_0\}$, and Perfect Substitutes subjects maximize $U_i^{PS}(\pi_s,\pi_o)=\pi_s+\theta_i^{PS}\pi_o$.[4] We aim to estimate the fractions of S, PS, and PC subjects in the population, and how this varies with subject characteristics, which in this case is an indicator for sex (1 = female).

We now place some restrictions on the distribution of the nuisance parameters. Firstly, we assume that they can only take on positive values. Secondly, identification

[4]While the nuisance parameter in the non-selfish types enters the utility functions on different agents' payoffs, we choose this specification because increasing θ_i^τ results in more generous behavior for both types.

requires that one type cannot be parameterized by another. Without further restrictions, predictions of the non-selfish types approach the selfish predictions as $\theta_i^\tau \to 0$. We therefore require that the supports of the nuisance parameters do not include values close to zero. We therefore choose the following transformation to model the distributions of nuisance parameters:

$$\log(\theta_i^\tau - \delta^\tau) = \tilde{\theta}_i^\tau \sim \text{Normal Mixture}(\psi^\tau, \mu^\tau, \nu^\tau) \tag{2.43}$$

That is, conditional on the component.$\theta_i^\tau - \delta^\tau$ is log-normally distributed. We choose $\delta^{PC} = 1/9$, $\delta^{PS} = \frac{1}{3}$, which ensures that the PC types will pass at least 10% of the endowment when $p_k = 1$ (more for $p_k < 1$, less for $p_k > 1$), and that the PS types will pass their entire endowment in at least one treatment.

In order for our model to respect the fact that subjects' answers may be noisy, and are censored at whole numbers of tokens, we use the following latent choice rule:

$$t_{i,k}^* = \arg\max_{t\in[0,1]} U_i^{\tau_i}(t, (1-t)/p_k; \theta_i^\tau) + \epsilon_{i,k}, \quad \epsilon_{i,k} \sim \mathcal{N}(0, \gamma) \tag{2.44}$$

$$t_{i,k} = \min\{\max\{\text{round}(w_k t_{i,k}^*, 1)/w_k, 0\}, 1\} \tag{2.45}$$

where w_k is the endowment of tokens in treatment k.[5]

Table 2.4 reports posterior means and standard deviations of the multinomial probit parameters, however we encourage the reader to focus on Table 2.5, which shows the mixing probabilities by sex implied by the parameters estimated in Table 2.4. In particular, the final row of Table 2.5 shows a hypothesis test that females are more likely to be each type than males. For the rightmost column, we see that at all reasonable levels of significance males and females are no more or less liekly to be the selfish type; however in the non-selfish types we see statistically and economically significant differences. Firstly, women are more likely than men to be the PC type

[5]In their Table IV, Andreoni and Vesterlund (2001) introduce a fourth type (with nuisance parameters) and account for censoring at the minimum (0) and maximum (w_k) number of tokens. We account for censoring at the endpoints as well as censoring at integers.

Table 2.4.: Posterior moments of multinomial probit parameters β and Σ.

	PC	PS
COEFFICIENTS		
constant	-0.5279	-0.5469
	(0.2880)*	(0.2148)*
female (D)	0.7231	-0.4246
	(0.3791)*	(0.3205)
COVARIANCE		
PC	2.4724	0.806
	(2.4204)***	(0.6206)
PS	0.806	1.0
	(0.6206)	(0.0)***

Table 2.5.: Posterior means (standard deviations) of mixing probabilities. The bottom row shows the posterior probability that females are more likely to be this type than men.

	PC	PS	S
Male	0.3213	0.2176	0.4611
	(0.0022)	(0.0019)	(0.0026)
Female	0.5300	0.0925	0.3775
	(0.0046)	(0.0017)	(0.0045)
$\Pr[\tau\|f] > \Pr[\tau\|m]$	0.9952	0.0203	0.1597

(53.3% vs. 32.0% respectively), and women are less likely than men to be the PS type (9.2% vs. 21.8% respectively).

We now move to analyzing our model's fit to the experimental data. Figure 2.1 overlays Figure I from Andreoni and Vesterlund (2001) with the posterior mean demand predictions conditional on type.[6] Panel (a) shows the prediction conditional on being the PC type, and panel (b) shows the prediction conditional on being PS. Solid lines show the posterior means of the estimated model, and dashed lines show raw means from the experiment. At all prices these lines fit within their relevant 95% Bayesian credible region (dotted lines). Figure 2.2 shows posterior mean predictions of demand conditional on sex. For female subects (panel (a)), the raw data fall within the 95% credible region of the structural model at all prices, for males, the model fits within this band except when $p = 1$. At this point, we are able to revisit Andreoni and Vesterlund (2001)'s conclusion that "when altruism is expensive, women are kinder, but when it is cheap, men are more altruistic." We test this with our structural model by assigning a posterior probability that the mean demand curves shown in Figure 2.2 cross. With a posterior probability of about 92%, we interpret this as mild support for their claim.

For the sake of exposition, in this section we introduced some unobservable heterogeneity in the form of nuisance parameters θ_i^{PC} and θ_i^{PS}. One question this raises is how much heterogeneity in preferences exists within the types, and how restrictive the original assumption of $\theta_i^{\tau} = 1$ actually was. Figure 2.3 shows the posterior cumulative densities of the mean and standard deviations of the nuisance parameters. The left panel shows that the posterior means fall reasonably close to the assumed values with high probability, indicating that *on average* this assumption may not be so bad. However we cannot reject hypotheses that the means of these distributions are equal to one.[7] The right panel of Figure 2.3 shows very little probability mass close to zero,

[6]Andreoni and Vesterlund (2001) define "demand" as the other's monetary payoff divided by value of the endowment to the self. In our notation, this is equal to $(1 - t_{i,k}^*)/p_k$.

[7]95% Bayesian credible regions of $E(\theta_i^{PC})$ and $E(\theta_i^{PS})$ are [1.0679, 1.1911] and [1.1788, 1.5766] respectively. Both do not span the originally assumed value of 1.

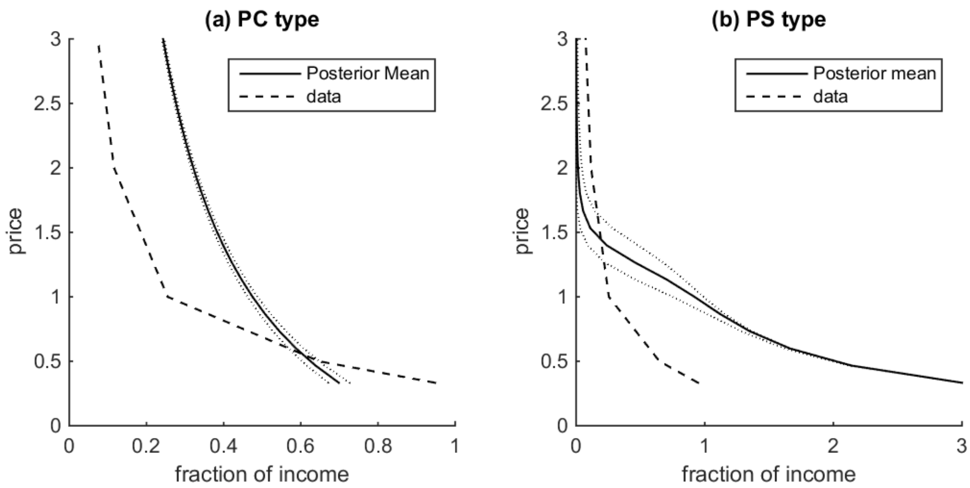


Fig. 2.1. Demand for other's income by type. Solid lines show posterior mean predictions, and dotted lines show a 95% Bayesian credible region. Dashed lines show the raw means of choices in the experiment. The selfish type's demand is omitted and falls on the vertical axis.

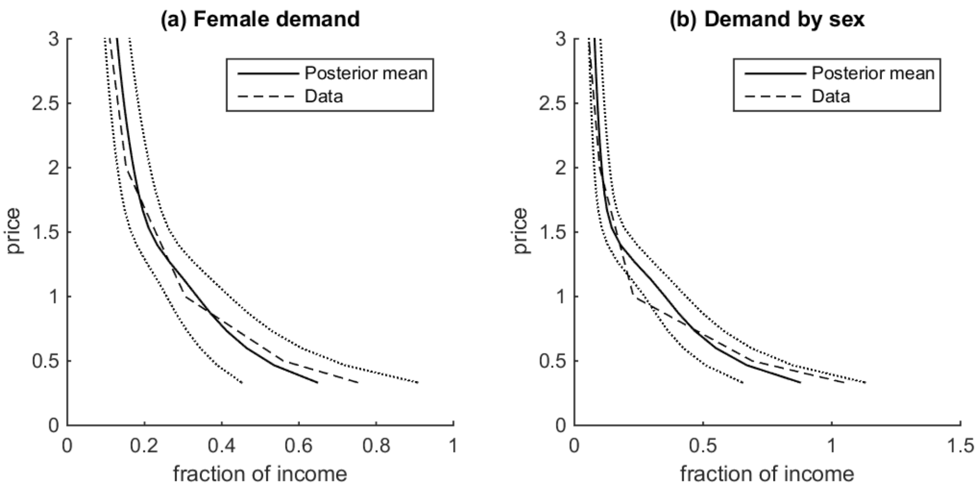


Fig. 2.2. Demand for other's income by sex. Solid lines show posterior mean predictions, and dotted lines show a 95% Bayesian credible region. Dashed lines show the raw means of choices in the experiment..

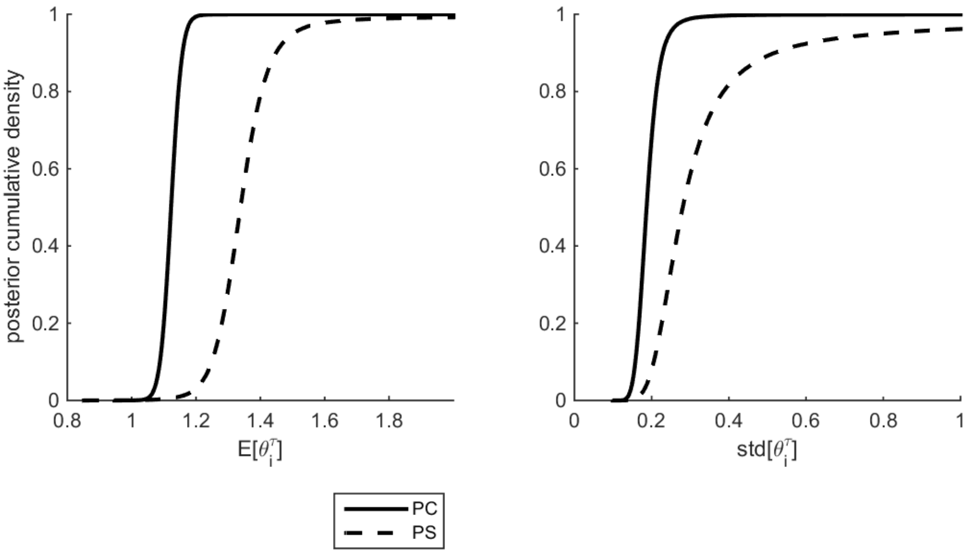


Fig. 2.3. Posterior distributions (cdf) of mean and standard deviation of the nuisance parameters.

and we can place 95% lower bounds on the standard deviations of $\text{std}[\theta_i^{PC}] > 0.1507$ and $\text{std}[\theta_i^{PS}] > 0.1907$. As these lower bounds are in the order of 10% of the assumed value, we interpret this as evidence in favor of economically significant heterogeneity: We reject Andreoni and Vesterlund (2001)'s assumption of $\theta_i^{\tau} = 1$ in favor of a heterogeneous θ_i^{τ}, and reject a weaker assumption that the restriction holds in expectation. It is important to note that our estimation restricts the non-selfish types to have sufficiently large nuisance parameters. What the econometrician cannot distinguish here is whether there are some (say) PC types with θ_i^{PC} close to zero. These types would be classified as selfish in our estimation. It could therefore be the case that the true mean of θ_i^{PC} is closer to zero, as here we are reporting the conditional mean $E[\theta_i^{\tau} \mid \theta_i^{\tau} > \delta^{\tau}]$, rather than the unconditional mean $E[\theta_i^{\tau}]$.

One may question whether this significant heterogeneity in *preferences* propagates into significant heterogeneity in *decisions* within each type. In order to shed light on this, we compute expected prediction errors for the PC type, whose optimal choice is:

$$t_{i,k}^{PC*} = \frac{\theta_i^{PC}}{p_k + \theta_i^{PC}}$$

with elasticity with respect to parameter θ_i^{PC}:

$$\frac{\partial t_{i,k}^{PC*}}{\partial \theta_i^{PC}} \frac{\theta_i^{PC}}{t_{i,k}^{PC*}} = \frac{p_k}{p_k + \theta_i^{PC}}$$

so for the "standard" dictator game parameterization of $p_k = 1$, assuming $\theta_i^{PC} = 1$ would lead to an absolute percentage error prediction in the fraction of tokens kept by the self of roughly: $\left|\frac{1}{1+1} \times (\theta_i^{PC} - 1) \times 100\right|$,[8] or half the percentage error in θ_i^{PC}. We use the distribution of θ_i^{PC} to estimate this prediction error at all prices studied in this experiment. These errors are shown in Figure 2.4, which show moments of the percentage prediction error associated with assuming the all PC types have $\theta_i^{PC} = 1$.

[8] We use this expression for exposition only. In Figure 2.4 we do not use this linear approximation, and opt for the exact value instead.

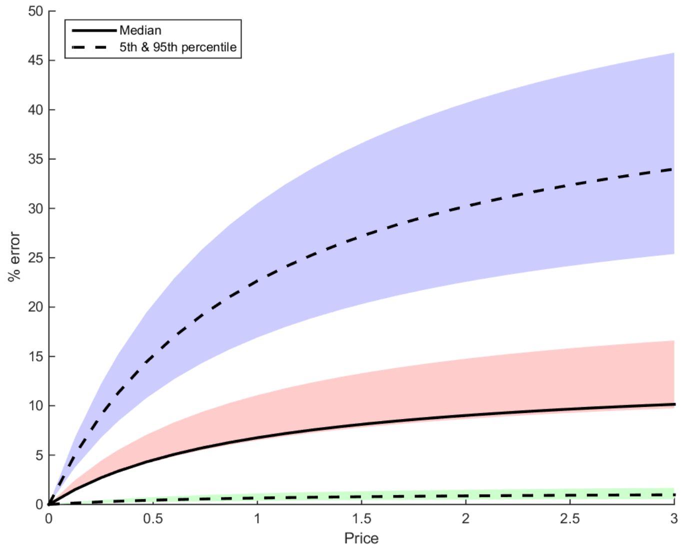


Fig. 2.4. Prediction errors associated with assuming $\theta_i^{PC} = 1$. The vertical axis shows the estimated percentage error in predicting the fraction of tokens kept by the self (i.e. $t_{i,k}^{PC*}$). The solid line shows the expected prediction error, and the dashed lines show the 5th and 95th percentiles of the prediction errors. That is, we estimate that 90% of individual prediction errors will fall between the dashed lines. The solid and dashed lines show posterior means, and the shaded regions are 95% Bayesian credible regions.

The expected prediction error is greater than 5% at prices equal to and greater than 1 (solid line) and at the highest price studied, $p_k = 3$, this prediction error is about 11% on average, and between about 29% for subjects in the 95th percentile of prediction errors.[9]

If the goal of an estimator is to assign a type for each subject, errors in predicted decisions are secondary to errors in how each subject is classified. We find that while many subjects' nuisance parameters may vary significantly from unity, assigning subjects to types based on the greatest posterior mean probability from our estimation produces remarkably similar results to the Andreoni and Vesterlund (2001)'s mini-

[9]Here we report posterior means of these moments. 95% Bayesian credible regions around these are [8.3%, 14.1%] and [21.1%, 39.7%] respectively.

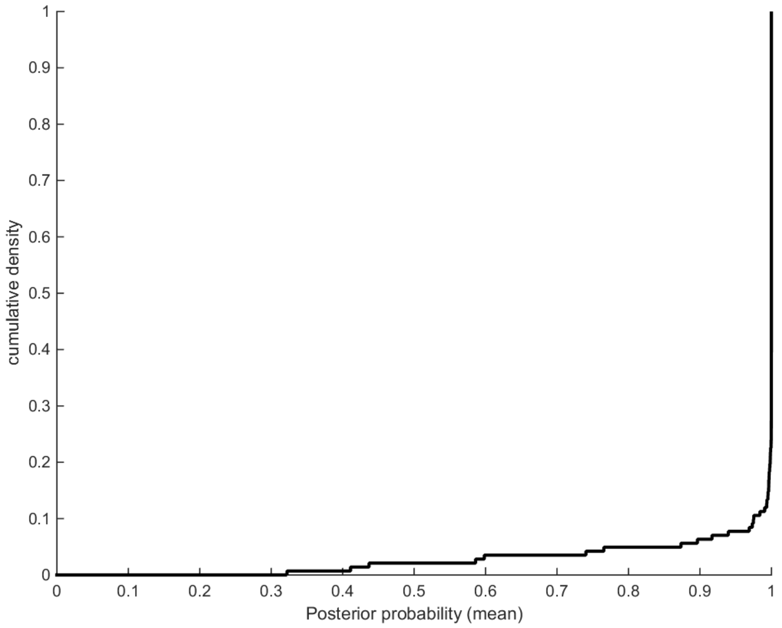


Fig. 2.5. Cumulative density of the mean posterior probability of each subject being the type assigned by Andreoni and Vesterlund (2001)'s minimum distance estimator.

mum distance method. Figure 2.5 shows the empirical cumulative density function of the probability that a subject is assigned to the same type as originally classified by Andreoni and Vesterlund (2001). If our classification were identical to their minimum distance method, the line would represent a degenerate distribution with a mass point at 1 (i.e. a horizontally-mirrored "L" shape). While there are some discrepancies, 93% subjects are assigned a posterior mean probability of at least 90% being the type that Andreoni and Vesterlund (2001) assigned them. In terms of a classification exercise, our method produces almost identical results.

Finally, while the individual nuisance parameters are not the focus of our estimation, we conclude our analysis of Andreoni and Vesterlund (2001) by showing shrinkage estimates of these for each subject in Figure 2.6. For the PC type, we see that almost all subjects are classified into either (1) not PC, or (2) PC. This

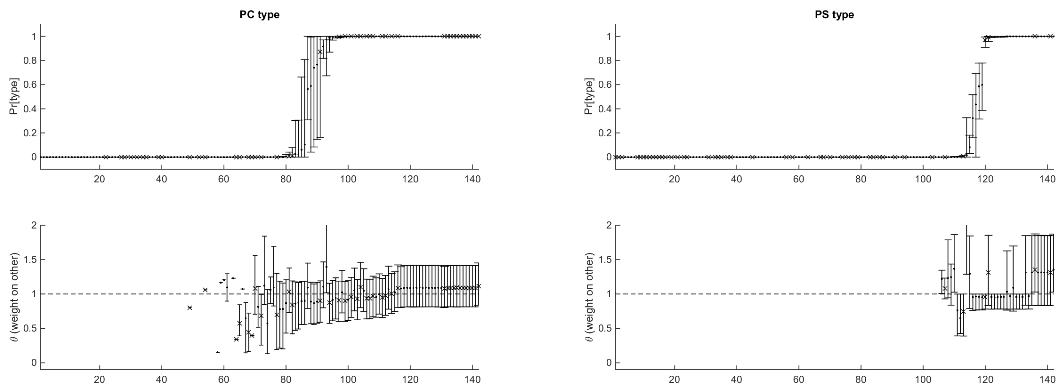


Fig. 2.6. Shrinkage estimates. The top row shows the posterior probability of being each non-selfish type, and the bottom row shows the nuisance parameter estimate for each non-selfish type. Error bars show a 95% Bayesian credible region around the point estimate, which is the posterior mean. Plots in each column are sorted by posterior mean probability of being that type. Estimates corresponding to female subjects are denoted with a "×" at the posterior mean.

is indicated by the very few points in the top-left panel with noticeable error bars. Classification into either PS or not PS is less precise. In the bottom panels of this figure we see estimates of the nuisance parameters. These become more relevant the further to the right of the plot we go, because these are the individuals more likely to be that type. Of interest here is that for many subjects the estimated parameter is reasonably far away from one, although at the individual level we frequently cannot reject that these parameters are equal to one.

2.6.2 Harrison et al. (2010)

Harrison et al. (2010) investigates discounting behavior and whether smokers discount differently to non-smokers. Their experiment includes four risk aversion tasks and six discount rate tasks. Each task was a series of Holt et al. (2002)-style binary choices. For example, one choice in a risk aversion task was between a "Lottery A" that paid 2000 DKK with probability 0.7 or 1600 DKK otherwise, and a "Lottery B" which paid 3850 DKK with probability 0.7 or 100 DKK otherwise. One choice in each of the four risk aversion task was selected for payment. Other choices in the risk aversion tasks varied across both probabilities and payoffs. One choice in the a discount rate task was between a "Payment option A" which paid 3000 DKK one month after the experiment, and a "Payment option B" which paid 3308 DKK seven months after the experiment. Other choices in the discount rate tasks varied the time before the first payment, and the payoffs. Subjects typically made 10 decisions in each of the ten tasks. A full description of the experiment design and procedure can be found in Harrison et al. (2005).

Their study asks two broad questions: Are smokers' preferences significantly different from non-smokers, and are smokers more or less likely to discount hyperbolically than non-smokers. We focus on this second question, which they answer with a Harrison and Rutström (2009)-style mixture model. These questions are useful to policy makers because knowing how or if smokers' preferences and behavior differ from non-smokers could help identify policies for encouraging smokers to quit, or discouraging non-smokers from starting.

They consider two models of discounting: exponential and hyperbolic. Subjects that discount exponentially choose the option that maximizes:

$$E\left[\sum_{t=0}^{\infty}(1+\delta_i)^{-t}\frac{x_t^{1-r_i}}{1-r_i}\right] \tag{2.46}$$

where x_t is the monetary payoff at time t, δ_i is i's discount rate, and r_i is i's CRRA risk aversion parameter. The expectation is taken at $t=0$. Hyperbolic discounters choose the option that maximizes:

$$E\left[\sum_{t=0}^{\infty}(1+\gamma_i t)^{-1}\frac{x_t^{1-r_i}}{1-r_i}\right] \tag{2.47}$$

where γ_i is i's hyperbolic discount rate. Importantly, note that an exponential subject's discount factor is $(1+\delta_i)^{-t}$, and a hyperbolic's is $(1+\gamma_i t)^{-1}$. Therefore at $t = 0$ when deciding about payments to receive at $t = 1$ and 2, an exponential discounter would never wish to revise their decision if given the opportunity at $t = 1$ (assuming no uncertainty), but a hyperbolic subject may wish to. The hyperbolic type is therefore considered the irrational, "behavioral" type. When trading off payments between $t = 0$ and $t = 1$, they types make identical decisions.

Before applying our estimator to their data, we note that for small discount rates, exponential and hyperbolic subjects make almost the same decisions.[10] This is demonstrated in Figure 2.7, where we simulate choices made by an exponential and a hyperbolic subject with the same risk preferences and discount rate. We shut down any decision errors here. Each line on this plot represents a different level of risk aversion (r_i), and the horizontal axis shows the discount rate. The height of each line is the fraction of choices that differ. Note that for small δ there is no difference between the two types. We therefore place a restriction the hyperbolic type's discount rate: $\gamma_i > 0.12$.

We apply our estimator to their data, using Harrison et al.'s specification of utility functions and decision error, but our specification of the nuisance parameters.[11] This is shown in Table 2.6. The first panel of this table shows the Probit parameters for the mixing probabilities. The dependent variable in this process is an indicator for hyperbolic discounting, therefore larger coefficients represent larger probabilities of being hyperbolic. Here we construct the explanatory variables without a constant, so the normal cdf of the probit parameter is directly interpretable as a mixing probability. This transform is shown in the second panel. Both smokers are non-smokers are

[10] The intuition behind this is as follows: For $\delta_i = 0$, both types will always choose the option that leads to the greatest amount of money, regardless of when the payment is made. For small δ_i and for the values of t studied in this experiment, $(1+\delta_i)^{-t} \approx (1+\delta_i t)^{-1}$, and so there is little difference between exponential and hyperbolic types' choices.

[11] Another difference with our estimation is that we assume that *subjects* are one type or another, but Harrison et al. (2010) assume types are determined at the *decision* level. We stick with the "once one type, always that type" specification in this paper, as is used in (for example) Conte et al. (2011). The "decisions are types" specification, as used in Harrison and Rutström (2009), could be implemented here without any extension needed.

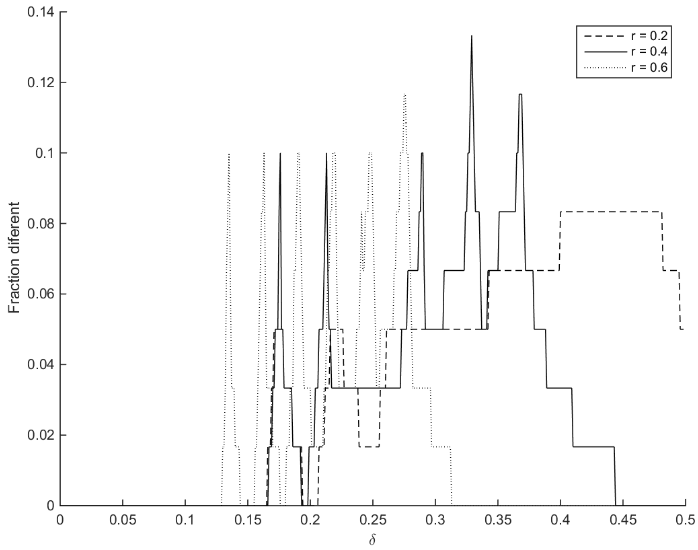


Fig. 2.7. Difference between decisions made by an exponential and hyperbolic discounter with the same level of risk aversion.

Table 2.6.: Mixture model estimation using data from Harrison et al. (2010), with smoking status used as an explanatory variable for the probability of being a hyperbolic discounter. The remaining fraction of subjects are exponential discounters.

	Mean	Std dev.		95% credible region	
Probit parameters					
smoker	-1.5503	(0.7373)		[-2.7622	0.1171]
non-smoker	-1.3751	(0.4866)	*	[-2.0126	-0.1142]
Mixing probabilities (transform of probit)					
smoker	0.1095	(0.1472)	***	[0.0029	0.5466]
non-smoker	0.1096	(0.1126)	***	[0.0221	0.4545]
log odds ratio	-0.4641	(1.0005)		[-2.9249	0.9570]
Exponential type distribution					
mean r	0.7124	(0.0530)	***	[0.5781	0.7736]
std r	0.3830	(0.4328)	***	[0.1661	1.3922]
mean $\log \delta$	-2.4542	(0.1137)	***	[-2.6431	-2.1709]
std $\log \delta$	2.9494	(4.5974)	***	[0.7843	13.3513]
$\text{corr}(r, \log \delta)$	0.0065	(0.1528)		[-0.3420	0.2682]
Hyperbolic type distribution					
mean r	0.4862	(0.1946)	*	[0.0867	0.7841]
std r	0.8014	(0.5997)	***	[0.2165	2.1774]
mean $\log \delta$	-2.3327	(0.2211)	***	[-2.9126	-1.9729]
std $\log \delta$	4.8273	(3.8255)	***	[1.5110	13.7021]
$\text{corr}(r, \log \delta)$	0.2691	(0.3855)		[-0.5756	0.9032]
Noise parameters					
μ	0.0652	(0.0054)	***	[0.0552	0.0744]
Error prob @ $u_A/u_B = 1.05$	0.3203	(0.0138)	***	[0.2924	0.3418]
ν	0.0052	(0.0007)	***	[0.0040	0.0065]
Error prob @ $u_A/u_B = 1.05$	0.0001	(0.0002)	***	[0.0000	0.0005]

estimated to by hyperbolic discounters with probability about 11%. These numbers are much smaller than the fractions estimated by Harrison et al. (2010). However this should not be surprising, as we do not allow subjects with discount rates below 0.12 to be classified as hyperbolic. One should therefore interpret this 11% as a lower bound on the fraction of hyperbolic discounters, as there could be more hyperbolic discounters present with discount rates below 0.12. Note that our credible region for the fraction of non-smokers who are hyperbolic spans 45% of the unit interval, while Harrison et al. (2010)'s interval spans just 18%: by this measure we can say less about a *lower bound* of the mixing probability than they claim to know about the probability itself. We then construct a 95% credible region around the log odds ratio of these mixing probabilities. As it contains zero there is little support that smokers are any more or less likely to be hyperbolic. Harrison et al. (2010) reach the same conclusion.

Harrison et al. (2010) also collect information on the number of cigarettes each participant smokes per day. In Table 2.7 we include this as an explanatory variable for the mixing probabilities, and find similar results to Table 2.6.

Subjects' individual probabilities of being hypobolic discounters are shown in Figure 2.8. We estimate approximately 70 subjects to be exponential with probability very close to 1. However this is an artifact of the restriction placed on the hyperbolic discount rate: as it is bounded away from zero, any subject with a small discount rate is dogmatically labeled an exponential discounter. As we move further right in this plot, subjects are more likely to hyperbolic, but the credible regions span most of the unit interval: the experiment does not provide much information about type.[12]

The experiment is, however, informative about discount rate and risk aversion. The individual-level estimates of these are shown in Figures 2.9 and 2.10 respectively.[13]

[12]Performing our estimation without the lower bound on the hyperbolic discount rate produces large credible regions for almost all of the subjects.

[13]As the discount rates of the two types are directly comparable for a decision involving only money now or in one unit of time, we construct a blended discount rate. Subjects estimated more likely to be hyperbolic receive more weight on their hyperbolic discount rate than their exponential.

Table 2.7.: Mixture model estimation using data from Harrison et al. (2010), with smoking status and number of cigarettes smoked per day used as explanatory variables for the probability of being a hyperbolic discounter. The remaining fraction of subjects are exponential discounters.

	Mean	Std dev.		95% credible region	
PROBIT PARAMETERS					
smoker	-0.1942	(1.0115)		[-2.1227	1.8858]
non-smoker	-1.3603	(0.4999)	*	[-1.9872	-0.0529]
ncigs	-0.5403	(0.6897)		[-2.1147	0.6538]
ncigs2	-0.0252	(0.0521)		[-0.1456	0.0491]
EXPONENTIAL TYPE DISTRIBUTION					
mean r	0.7159	(0.0524)	***	[0.5783	0.7749]
std r	0.3952	(0.5013)	***	[0.1657	1.5745]
mean $\log \delta$	-2.4570	(0.1187)	***	[-2.6433	-2.1517]
std $\log \delta$	3.4099	(5.7885)	***	[0.7936	17.6935]
$\text{corr}(r, \log \delta)$	-0.0063	(0.1549)		[-0.3687	0.2682]
HYPERBOLIC TYPE DISTRIBUTION					
mean r	0.4667	(0.1988)	*	[0.0580	0.7743]
std r	0.8191	(0.6744)	***	[0.2297	2.2384]
mean $\log \delta$	-2.3181	(0.2033)	***	[-2.8156	-1.9679]
std $\log \delta$	5.0798	(4.6735)	***	[1.4891	14.3115]
$\text{corr}(r, \log \delta)$	0.2761	(0.3917)		[-0.5852	0.9102]
NOISE PARAMETERS					
μ	0.0649	(0.0057)	***	[0.0552	0.0752]
Error prob @ $u_A/u_B = 1.05$	0.3192	(0.0145)	***	[0.2922	0.3432]
ν	0.0052	(0.0008)	***	[0.0040	0.0066]
Error prob @ $u_A/u_B = 1.05$	0.0001	(0.0002)	***	[0.0000	0.0006]

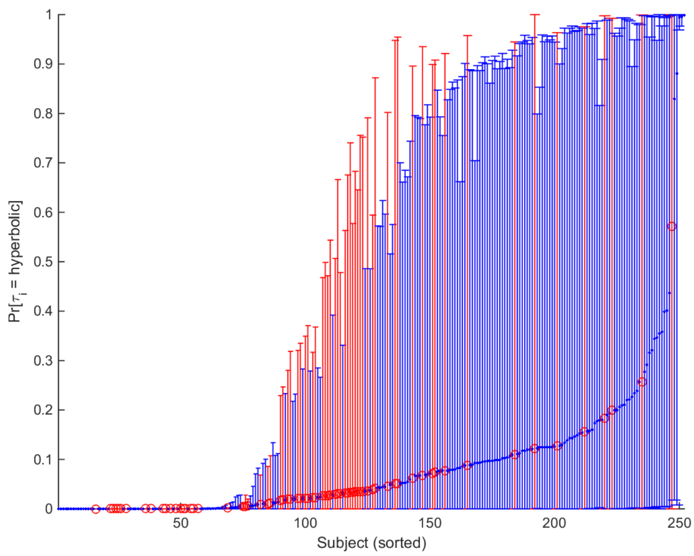


Fig. 2.8. Posterior mean probability of subjects being the hyperbolic type. Error bars show a 95% Bayesian credible region of this probability. Smokers are denoted by a "○" at the posterior mean. These estimates correspond to the specification reported in Table 2.7. .

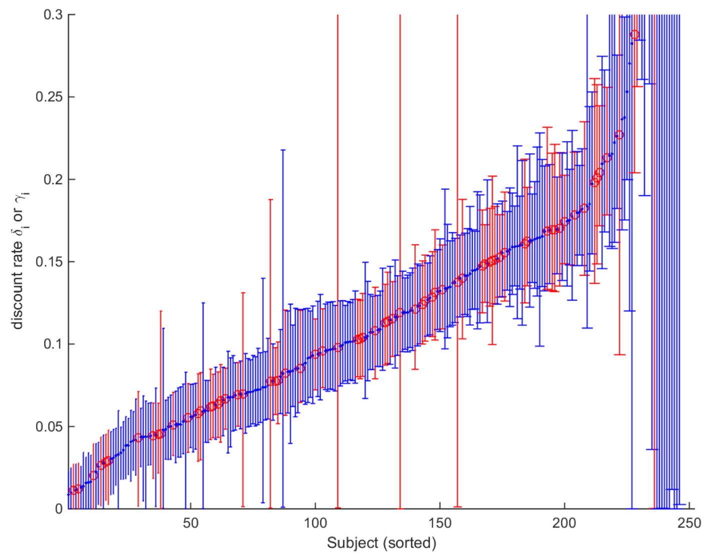


Fig. 2.9. Posterior means of subjects' discount rates. Error bars show a 95% Bayesian credible region for this value. Smokers are denoted by a "∘" at the posterior mean. These estimates correspond to the specification reported in Table 2.7.

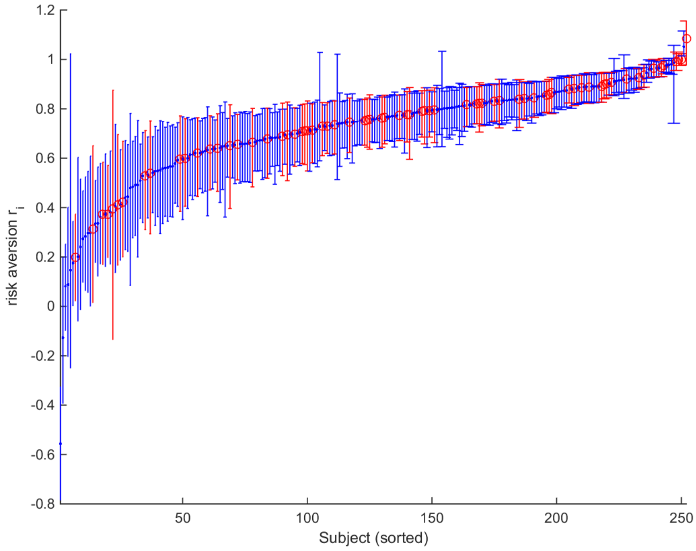


Fig. 2.10. Posterior means of subjects' risk aversion parameters. Error bars show a 95% Bayesian credible region for this value. Smokers are denoted by a "○" at the posterior mean. These estimates correspond to the specification reported in Table 2.7.

2.7 Conclusion

When considering alternative models of decision-making, we are usually more interested in their prevalence in our sample than the particular parameters for each subject that enter into each model. An individual is unlikely to be made better off by being forced to behave as if they are less risk averse than they would otherwise choose on their own, but could be made better off by being more Bayesian or less Prospect Theory. Estimating the prevalence of different decision rules is therefore of great economic significance, and motivates the use of mixture models. One complication of mixture models is that econometricians must take a stand on how nuisance parameters enter into their specification. We show that if this part of the model is incorrectly specified, the errors can propagate through to errors in the estimates of mixing probabilities. This can lead the researcher to falsely conclude that there are significant fractions of both (maybe more than two) types in the population, even when in reality only one decision rule is used.

We propose a technique that places less structure on the nuisance parameters, but still allows for estimation of the mixing probabilities, and how they vary by subject's characteristics. Our estimator inherits some desirable properties of existing approaches: mixing probabilities can vary by subject characteristics (Harrison and Rutström, 2009), nuisance parameters enter the model as random effects (Conte et al., 2011), and we can accommodate more than two types (Andreoni and Vesterlund, 2001). If the researcher also wishes to make statements about how nuisance parameters vary with subject characteristics, an extension of our work could use the $\theta_i^\tau = X_i\delta^\tau + \epsilon_i^\tau$ specification mentioned in the literature review. We omitted this from our analysis due to computational power issues, but the additional coding required is minimal within the Bayesian framework. This extension could be particularly useful in for example Harrison et al. (2010), if instead of mixing probabilities, we were primarily interested in risk aversion and time preferences. In this case the mixture model could help us make statements about these while remaining somewhat agnostic about the decision rule subjects use. Additionally, this could be used to either validate or reject Harrison and Rutström (2009)'s econometric simplification that observable characteristics uniquely determine nuisance parameters.

We analyze two experimental datasets which lend themselves to estimating a mixture model. We are unable to pin down the prevalence of hyperbolic discounting in our sample as precisely as Harrison et al. (2010) claim in their analysis. Deeper inspection of their experiment design reveals that while the experiment is very informative about subjects' risk and time *preferences*, it provides little information about discounting *behavior*. In our estimation this manifests itself as wide distributions of the posterior mixing probabilities. Since Harrison et al. (2010) make a dogmatic assumption about the role of unobservable nuisance parameter heterogeneity for observationally equivalent subjects (i.e. there is none), their model must accommodate the data with the limited tools at its disposal: any difference in choices of two subjects with the same observable characteristics *must* be attributable to different decision rules. There is no scope for different nuisance parameters.

We also revisit Andreoni and Vesterlund (2001), and test whether their conclusion that men and women are differently altruistic can be motivated structurally. While our estimation casts doubt on their assumption about nuisance parameters, we find as they do that non-selfish men and disproportionately likely to have perfect substitutes preferences compared to women. This structurally supports their result that men's demand for altruism is more price sensitive. We also generally agree with their classification of individual subjects into the three types.

Overall, this paper highlights some issues relevant to both experimental economists and econometricians. The results from even a correctly specified mixture model are only as reliable as the power of the experiment to distinguish between types. A good experiment design will anticipate this. Even if the experiment design is good, as econometricians we must worry about nuisance parameters: strong assumptions can lead to false conclusions about the mixing probabilities. When a mixture model is considered in the design phase, experimenters should think about how the experiment will separate types for all possible parameterizations, or at least for a plausible range of economically significant values. When analyzing the output from a mixture model, one should be aware of the types and their parameterizations that are distinguishable, and aim to make statements about only those that are.

3. HOSPITAL-INSURER BARGAINING POWER AND NEGOTIATED RATES

with Amanda Cook

3.1 Introduction

Health insurance in the US performs two functions for consumers. The first is the traditional function of transferring wealth from good to bad states of the world. The other is the result of the increased bargaining power that follows from having an insurance company act (in part) on behalf of a large number of customers. Insurers can, and do, negotiate prices with hospitals so that, for the same services provided, the hospital is paid less for the insured patients than it would otherwise charge an uninsured patient. The benefit to insurers of this ex-ante negotiation is that their costs are lower. Some of this discount may also be passed on to consumers in the form of lower premiums.[1] For hospitals, negotiation can ensure higher volumes of admissions, because insurers can direct their customers away from hospitals that fail to reach an agreement, typically by having a higher co-insurance rate for "out-of-network" hospitals compared to those that are in network.

Using data from Massachusetts Center for Health Information Analysis (CHIA), we estimate these negotiated rates from patient-level insurance claims and hospital discharge data. We find considerable variation in the amount paid for (normalized) services. This variation comes from three sources. (i) At any given hospital, the payment made for services depends on the patient's insurer. This variation can be interpreted as insurers having different bargaining powers with the hospital. (ii) Hold-

[1] Once insurance companies have negotiated a lower payment. They may either rent-seek or pass the savings on to consumers. While it is not clear which of these actions the insurance company will take, from the point of view of social welfare, we would be interested in the latter.

ing the insurer constant, the payment for services depends on the hospital at which the patient was treated. We interpret this as evidence of hospitals having different bargaining power over insurers. (iii) Considerable variation in these payments is left over, even after controlling for the patient's hospital and insurer. The existence of this residual variation is testimony to the imperfectly-competitive, and imperfectly understood nature, of health care markets.

The contribution of this paper is that we (1) calculate variation within a hospital-insurance pair, (2) examine how this within-pair variation compares to the variation of prices between different hospital-insurer pairs. In doing so, we contribute to the literature studying competition in health care and insurance markets, by adding a patient-specific level of observation whereas previous studies such as Cooper et al. (2015) and Ho and Lee (2015) have used a hospital as the level of observation.

3.2 Literature Review

Features of hospital-insurer negotiations: Gaynor et al. (2015) outline four elements that characterize negotiation between hospitals and insurers, and the incentives faced by insured individuals:

1. Insurers can encourage patients to go to particular hospitals over others through differential co-insurance rates. Hospitals with which the insurer has negotiated a favorable price have a low "in-network" rate. Hospitals with which the insurer has not negotiated have a higher "out-of-network" rate.

2. Except for co-insurance, patients do not pay directly for their care.[2]

3. At the time inpatient care is needed, the patient is generally locked-in to their insurance plan, and so cannot change plans to access the negotiated rate of another insurance product that would (ex-post) be better to treat their illness.

[2]While deductibles are common in the US health insurance industry, it is not the focus of Gaynor's study, nor does it affect hospital-insurance negotiations as it is focused on the spitting of the payment between insured individuals and insurance companies.

4. Insurers negotiate with hospitals over network status (patients pay lower co-insurance rates at in-network hospitals),and how the hospital will be compensated for services. In our dataset, the reimbursement structure is mostly on a fee-for-service basis (e.g.: the insurer pays a fraction of the amount charged by the hospital), or a DRG basis (e.g.: the insurer pays the hospital a flat rate determined by the patient's diagnosis and severity of illness).

Reinhardt (2006) documents some of the features of how payments are made between hospitals and insurers. Each hospital has a "charge master", a list of prices for each procedure performed in the hospital. Charges reported by the hospital are the sum of the prices from this list of all services performed on a patient. For insured patients, charges typically differ from the amount the hospital is compensated (by the insurer and patient together), Compensation is based on a negotiated rate agreed upon between the hospital and insurer. An implication of this, and the focus of our study, is that a hospital's compensation for a particular patient depends on the patient's insurer, and by extension the relative bargaining power of the hospital and the insurer during the negotiation.

Market power: Gaynor et al. (2015) document significant and increasing concentration of both hospitals and insurers. Data on insurers suggests that they are not price-takers either when dealing with employers, or with health care providers. Dafny (2008) provides evidence that insurers have market power over employers in that they can extract some surplus from more profitable firms through higher premiums. If the insurance market were competitive, premiums for the same health plan would be the same across firms. Dafny et al. (2009) also show that premiums increase when insurers merge. On the provider side of insurers' operations, they also find a decline in physicians' compensation and employment when insurers merge. There is also evidence of varying negotiating power on the hospital side of the market. For example, Ho (2008) estimates that "star" hospitals and capacity-constrained hospitals can command markups of about 25% of revenues. Other estimated markups were much lower. Our analysis generally agrees with this markup: we estimate that academic medical centers with an emergency room are compensated approximately 23% more than community hospitals without an emergency room.

Cooper et al. (2015) use insurance claims data of individuals with private employer-sponsored insurance to document variation in prices paid to hospitals. They estimate prices paid for privately insured individuals that vary by a factor of three over the Hospital Referral Regions. These prices are not strongly correlated (0.14) with similar compensations for Medicare beneficiaries, but are associated with measures of market power, Herfindahl indexes and counts of hospitals within a market, suggesting that hospitals' compensation for the privately insured is not driven by cost. Prices for specific procedures vary by even greater factors. Our analysis complements that of Cooper et al. (2015) by exploiting a more detailed data set, albeit only including observations from Massachusetts.

If variation in negotiated prices is driven by varying levels of competition, then prices should be lower in markets with more competitive hospitals. Ho and Lee (2015) use exposure to Kaiser Permanente, a large vertically integrated insurer, as a measure of competition in a health care market. Hospitals in more competitive markets typically have lower prices, however attractive hospitals are able to negotiate higher rates.

Welfare implications for insured individuals: Negotiated prices could affect patients' welfare for several reasons. Firstly, since insured patients typically pay a fixed fraction of the medical bill (the co-insurance rate), a lower negotiated rate means lower out-of-pocket expenses.This effect may be somewhat mitigated by hospitals' response to lower prices. Using an exogenous change in Medicare pricing, Dafny (2003) shows that hospitals respond to lower prices by "upcoding" patients to diagnosis codes that attract higher reimbursement, although she finds no evidence of an increase in the intensity of care. Secondly, depending on their market power, insurers may pass some of the savings on to consumers in the form of lower premiums. However Dafny et al. (2015) show that premiums are typically lower in markets with more insurers. At least by this measure, it seems that the between-insurer competition effect on premiums is stronger than any passed-through costs of higher negotiated prices in more competitive markets.[3] Ho and Lee (2015) find consolidation of insurers increases premiums and find a heterogeneous effect on negotiated prices. Another concern for the welfare of consumers could be through the restricted choice set imposed by an insurer's network. Assuming no price effect, Ho (2006) estimates a welfare loss of about $1 billion per year in the markets studied due to restricted choice sets. As Ho notes, this loss may be mitigated by price reductions at the hospitals remaining in the insurer's network. Using a natural experiment, Gruber and McKnight (2016) find that consumers are highly price sensitive to limited-network insurance plans, and that those who switch to them had almost 40% lower medical expenses than those who did not. This reduction in expenditure is attributable to reductions in both demand and price for services, especially for specialist and hospital care. They found substitution effects in the direction of primary care.

3.3 Data Description

We use two datasets produced by the Massachusetts Center for Health Information and Analysis (CHIA). The first of these is the All Payer Claims Database (henceforth APCD), the set of records submitted by insurance companies to the state of Massachusetts, and provides detailed information about charges, the amount paid by the insurance company, the patient's out-of-pocket amount, details on deductibles. The second is the Massachusetts Acute Hospital Case Mix Database (henceforth Case Mix), the set of medical records submitted by the medical providers to the state. While the data set is rich, including outpatient care, we restrict ourselves for this

[3]If we make the assumption that patient illnesses are spread evenly across insurance companies, then the actuarially fair price of insurance should be constant. If this were the case, one could compare premiums to determine measures of surplus. In a perfect world we would know both actuarially fair prices and premiums, and look at the difference between the two.

study to inpatient hospital services only. These inpatient records include demographic information as well as detailed information about diagnosis, procedures, etc.

We capitalize on the rich information about payments (in APCD) and patients (Case Mix) through an aggregated merge of the two datasets. The inpatient records in both Case Mix and APCD are a census of hospital admissions of privately insured patients for the state of Massachusetts. The two datasets are distributed yearly, but since reporting periods differ overlap for nine months of the year. Therefore we observe every admission of a privately insured patient in Massachusetts between January and September of 2013 (inclusive). To this end, we first collapse the APCD by patient episode (multiple line items are associated with each patient episode),[4] then use episode-specific characteristics documented in both datasets to merge the two files. These include patient age, gender, residential zip code, hospital zip code, and at least one of the following: diagnosis code, principal procedure code and total charges within \$1000. For the overlapping 9-month period of the two datasets, we match approximately 65% of privately insured patients. Of the patients we do match, over a third of them match on all three criteria (principal diagnosis, principal procedure, and total charges within \$1000), about 40% match on two criteria.

3.3.1 Some important variables

Table 1 shows summary statistics of some variables in the merged dataset. Summary statistics are calculated separately for two different payment methodologies, Fee For Service (FFS) and Diagnosis Related Group (DRG), which are discussed in more detail below. Of particular interest are the first three variables summarized in this table, `allowed` `charged` and `paid`, which capture different ways of accounting for hospital services.

[4]For example, a patient may go to the hospital and have separate charges from radiology and surgery. These would show up as separate line items in APCD. Since we are interested in all charges associated with a single patient episode, we sum charges from each line item for each patient episode.

Table 3.1.: Summary statistics

	(1) FFS		(2) DRG	
	mean	sd	mean	sd
allowed	6902	12481	10076	12645
charged	19171	21954	20281	23210
paid	6771	11064	10810	13956
weight_imp	1.126	1.124	1.369	1.484
aoverwi	6095	7456	8066	6719
poverwi	6051	7472	8992	8337
coverwi	16860	8320	15356	7684
Observations	47739		35855	

The first variable, `allowed`, documents the maximum amount to be paid to the hospital for a particular illness or procedure. Specifically, the CHIA Submission Guide instructs:

> Report the maximum amount contractually allowed, and that a carrier will pay to a provider for a particular procedure or service. This will vary by provider contract and most often it is less than or equal to the fee charged by the provider.
>
> APCD Medical Claims File Submission Guide, p 38.

If patients were homogeneous, the mean of the `allowed` variable would represent negotiated rates between the hospital and the insurance company.

The second variable `charged` is the amount the hospital charged for the services received by a patient for a particular visit. The documentation for this variable suggests it is independent of insurance type and payment methodology. One could view this as the 'sticker price' for services, or, following the language used in Reinhardt (2006), we could think of these as the 'charge master' prices. It is worth observing that the charges are very similar between the two payment methodologies: DRG and FFS.

The `paid` variable is the total amount paid to the hospital. There are, however, a few different methodologies to determine payments, the most frequent of which is fee-for-service. As the name suggests, charges are accrued for each service. For example, for each x-ray the insurance company pays $150. This methodology for payments accounts for about 40% of our inpatient records.

The next largest payment methodology in our data set is a Diagnosis Related Group (DRG) system. DRG systems pay based on the patient illness, not on services received. For example, a patient comes in and is admitted with pneumonia. Once the diagnosis is determined (in this example, pneumonia), the patient is assigned to a DRG and based on that DRG the insurance company pays the hospital a flat fee (say $7000). This payment is invariant if the patient receives a great deal of services or a very few services. DRG payments account for about 35% of our data. While there are a few other payment systems, no other single system accounts for a large fraction of our data. As these two payment methodologies are quite different, summary statistics are provided separately for both.

It is important to note that for some patients `paid` is larger than `charged`. At first, this is quite counter-intuitive. Why would one pay more than one is charged for services? However, it is precisely because of the variation in payment methodologies that the paid to charged ratio can be greater than 1. Take a patient who is admitted under a DRG system. This patient's payments are determined by the DRG classification, not by the services they receive. As such, their payment may exceed their charges, which is a function of the services they actually receive in the hospital.

As not all illnesses require the same level of service provision, a weighting system was developed by Medicare and Medicaid. These weights, which are based on illness classification, allow us to compare average service provision for numerous illnesses. We had DRG illness classifications for approximately 40% of our data.

The next variable `weight imp` ("Imputed Weight") is a derived variable equal to patient's charges relative to the casemix-adjusted average charges at the patient's hospital. The reason for this normalization is that we wish to attribute differences in our estimates of negotiated rates to differences in aspects of the bargaining process, rather than hospitals having different charge-master prices.[5] We interpret this variable as an aggregation of the quantity of services received, relative to the services a patient at that hospital would receive if they had a DRG weight of 1. Therefore, a patient with an imputed weight of 2 receives twice the services (measured in charges) than a patient with an imputed weight of 1. It is worth noting that the average weight of fee-for-service patients is lower than that for diagnosis related group patients. We would interpret this to mean that our DRG patients are 'sicker' and require about $1.37/1.13 - 1 \approx 21\%$ more services than out FFS patients.

[5] We motivate this normalization in more detail in Appendix C.1, and describe the procedure for calculating imputed weight in Appendix C.2.

We have also included the variables `aoverwi` (allowed over weight) and `poverwi` (paid over weight) to be able to better compare across payment types. When we look at an illness adjusted measure of allowable services, the two values $6,100 (for FFS) and $8,065 (for DRGs) are far closer than when we consider the raw means $6,933 (for FFS) and $10,104 (for DRGs). Similarly, when we compare `poverwi` (paid over weight), we see that the difference in payment for equally ill patients between FFS and DRG has dropped ($6052 vs $8991), compared to their per-patient averages without accounting for illness levels($6791 vs $10843). The goal of this paper is to understand the 'markdown' of charges for normalized patients.

3.4 Results

We begin this section by testing our assumption that FFS payments are a fraction of charges. Table 3.2 shows the result of regressing log payments against log(imputed weight) for FFS payments, with various controls. If FFS payments are truly a fraction of charges, then we would expect a coefficient of 1 on log(imputed weight) in a

Table 3.2.: OLS regressions with FFS log payments against log(charges), with various controls. Significance stars are suppressed. The row labeled "H0: fixed markdown" reports the p-value for the test that the coefficient on log(weigh imp) is equal to 1 (i.e., a linear relationship between payments and imputed weight).

	(1)	(2)	(3)	(4)	(5)	(6)
log(weight imp)	1.040	1.029	1.009	0.740	0.737	0.723
	(0.0106)	(0.0106)	(0.0104)	(0.00922)	(0.00909)	(0.00908)
N	47278	47278	47278	47278	47278	47278
R^2	0.168	0.182	0.226	0.417	0.441	0.467
AIC	191793.2	191004.3	188456.6	175050.1	173110.9	172188.6
BIC	191810.7	191056.9	188947.4	175260.4	173794.4	178630.0
H0: fixed markdown	0	0.00600	0.364	0	0	0
Controls						
EMS region	-	Y	-	-	-	-
Hospital	-	-	Y	-	Y	Y
Insurer	-	-	-	Y	Y	Y
Interactions	-	-	-	-	-	H $\times$ I
No. of controls	0	4	54	22	76	733

Standard errors in parentheses

regression of log(payments) on log(imputed weight) and control variables.[6] That is, letting c_i denote the controls used, the relationship would be:

$$\log(\text{paid}_i) = \log(\text{weight}_i) + \tilde{\rho} c_i \tag{3.1}$$

The row labeled "H0: fixed markdown" reports the p-value for this test. While we reject the hypothesis that the coefficient is 1 for all but model (3), we note that in models (1), (2), and (3) that the coefficients are close to one, indicating a close to linear relationship. Looking at the coefficient on log(weight_imp) in columns (4) (5) and (6), we see some variation between 0.723 and 0.740. Given our understanding of FFS payments, we have run the OLS regressions from columns 5 (which includes hospital and insurer fixed effects) and have plotted the coefficients for the general model and one in which we impose that the coefficient on log(imputed weight) is 1. These coefficient plots can be found in Appendix C.4. What we see in these plots is that while the slope coefficients differ significantly from 1, our estimates of negotiated rates do not vary much between these two specifications: the ranking and approximate levels are very close.

We also investigate the relationship between (log) imputed weight and (log) payments for patients under the DRG payment systems. Table 3.3 shows the result of regressing log(payments) against log(imputed weight) for DRG payments, with various controls. In the DRG system, the hospital is reimbursed for treating a patient with a particular disease, and this payment is independent of the services the patient actually receives. We observe that while the coefficient on log(imputed weight) is somewhat more consistent across the different specifications, that the predictive power of the models is lower. This is consistent with our understanding of how reimbursement works for patients under a DRG system.

[6] If we assume charges are a linear combination of services and charge master prices, then we expect insurance companies to pay a fixed fraction of these prices. In a log specification, then, for a 10% increase in charges we would expect a 10% increase in payments, which is why we assert that this coefficient should be equal to 1.

Table 3.3.: OLS regressions with DRG log payments against log(inputed weight), with various controls. Significance stars are suppressed. The row labeled "H0: fixed markdown" reports the p-value for the test that the coefficient on log(weigh imp) is equal to 1 (i.e., a linear relationship between payments and imputed weight).

	(1)	(2)	(3)	(4)	(5)	(6)
log(weight imp)	0.929	0.929	0.892	0.924	0.884	0.884
	(0.00987)	(0.00983)	(0.00930)	(0.00992)	(0.00934)	(0.00934)
N	35631	35631	35631	35631	35631	35631
R^2	0.199	0.206	0.301	0.200	0.303	0.309
AIC	136066.6	135761.5	131319.7	136049.1	131259.9	131168.5
BIC	136083.6	135812.4	131794.6	136116.9	131785.7	132635.7
H0: fixed markdown	0	0	0	0	0	0
Controls						
EMS region	-	Y	-	-	-	-
Hospital	-	-	Y	-	Y	Y
Insurer	-	-	-	Y	Y	Y
Interactions	-	-	-	-	-	H $\times$ I
No. of controls	0	4	54	6	60	171

Standard errors in parentheses

Next, we investigate the explanatory power of individual hospitals and insurers, and their interactions, in explaining the ratio of payments to charges.[7] This ratio is of interest because it represents the fraction of charges for which the hospital is compensated. In a fee-for-service environment, this ratio provides preliminary insight into the rate negotiated between hospitals and insurance companies. We look at a number of specifications using different controls, and evaluate the predictive power of different models using both frequentest and Bayesian criteria. We consider both linear and log specifications.

Table 3.4 reports various measures of goodness-of-fit for regressions of the paid to charge ratio against controls for emergency medical service (EMS) regions, hospitals, and insurers. Each column restricts the sample to a single payment type. Columns 1 and 2 report results when restricting the sample to fee-for-service (FFS) payments with linear and log specifications. Since FFS payments are typically negotiated as a fraction of charges, we would expect these regressions to have better fit than those with DRG payment types. Looking at Columns (3) and (4), we see the same regressions (both linear and log) for DRG payments. The better fit of the FFS model is observed in the systematically higher R^2 values reported in the first two columns, corresponding to FFS payments. We also observe that in both FFS and DRG payment methods, the log specification fits the data better than the linear counterpart. This is consistent with our understanding of how the industry handles reimbursement.

The bottom two panels report the Bayesian and Akaike information criteria for the estimated models. For FFS payments, the BIC selects Model 4, which controls for hospitals and insurers, but without interactions.[8] The AIC selects Model 7, which includes hospital-insurer interactions. This is due to the AIC and BIC differing in the penalty function for adding more variables. For the BIC it is $k \log(N)$, where k and N are the number of controls and the number of observations respectively. For the

[7]Looking at the ratio of payments to charges is a special case of regressing log(charges) against log(payments), in which we are forcing the coefficient on log(charges) to be 1.

[8]Note in Table 2 the value for Model 4 is the lowest. As the BIC is a penalty function, lower numbers indicate better performance.

Table 3.4.: Regressions of the ratio of payments to charges, controlling for EMS region, hospitals, insurers, and interactions thereof.

	(1) poverc	(2) log_poverc	(3) poverc	(4) log_poverc
Mean value	0.383***	-1.962***	0.659***	-1.129***
	(172.99)	(-234.16)	(212.78)	(-127.90)
N	47739	47739	35855	35855
Payment type	FFS	FFS	DRG	DRG
NUMBER OF CONTROLS				
1 - EMS region (E)	4	4	4	4
2 - Hospital (H)	61	61	61	61
3 - Insurer (I)	22	22	6	6
4 - H & I (uninteracted)	83	83	67	67
5 - H, I, & E (uninteracted)	85	85	68	68
6 - H & (I & E interacted)	154	154	84	84
7 - H & I (interacted)	779	779	179	179
R^2				
1 - EMS region (E)	0.00416	0.00682	0.00319	0.00436
2 - Hospital (H)	0.0565	0.0613	0.157	0.165
3 - Insurer (I)	0.139	0.258	0.00103	0.000808
4 - H & I (uninteracted)	0.191	0.310	0.158	0.166
5 - H, I, & E (uninteracted)	0.191	0.310	0.158	0.166
6 - H & (I & E interacted)	0.197	0.314	0.159	0.166
7 - H & I (interacted)	0.234	0.342	0.163	0.173
BAYESIAN INFORMATION CRITERION				
0 - Constant only	66093	193233	63456	138603
1 - EMS region (E)	65937	192950	63384	138488
2 - Hospital (H)	63974	190868	57981	132792
3 - Insurer (I)	59164	179220	63482	138637
4 - H & I (uninteracted)	56850	176394	58003	132801
5 - H, I, & E (uninteracted)	56866	176401	58012	132808
6 - H & (I & E interacted)	57286	176922	58149	132961
7 - H & I (interacted)	61786	181668	58971	133663
Model selected	4	4	2	2
AKAIKE INFORMATION CRITERION				
0 - Constant only	66093	193233	63456	138603
1 - EMS region (E)	65902	192915	63350	138454
2 - Hospital (H)	63438	190333	57463	132274
3 - Insurer (I)	58971	179027	63431	138586
4 - H & I (uninteracted)	56122	175666	57435	132233
5 - H, I, & E (uninteracted)	56120	175656	57435	132231
6 - H & (I & E interacted)	55934	175571	57436	132248
7 - H & I (interacted)	54951	174833	57452	132144
Model selected	7	7	4	7

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

AIC the penalty is $2k$. As N is relatively large in our sample, these two criteria place greatly different weight on the number of controls. We interpret these selections as hospitals and insurers having varying bargaining power, which is supported by both criteria. We interpret the AIC's selection as hospital-insurance pairs being the level of observation that best explains the variation in price.

For DRG payments, the BIC selects Model 2, which controls for hospitals. The AIC selects Model 4 with a linear specification for the ratio of payments to charges, which includes all hospitals and insurance companies without interactions. The AIC selects Model 7 in the log specification, which includes all interactions. We now turn our attention to understanding the variation in negotiated rates at the hospital-insurance level.

3.4.1 Estimates of negotiated rates for FFS payments

In this section, we estimate negotiated rates for FFS payments, and document the variation in rates across hospitals and insurers. We investigate negotiated rates using payments per imputed weight as the relevant metric. Figures 3.1 and 3.2 show coefficient plots from a regression of the (log) paid per imputed weight ratio against indicators for hospitals and insurers respectively, suppressing the constant term. Differences in these coefficients represent fractional negotiated rate differences. That is, if hospital A has a coefficient of 8.5, and hospital B has a coefficient of 9.0, then hospital B's negotiated rates are estimated to be on average $\exp(9.8 - 8.5) \approx 1.65$ times greater than hospital B's. These figures document much variation in mean compensation per weight by both hospital and insurer. For hospitals (Figure 3.1) we see variation in negotiated rates by a factor of approximately $\exp(10.4 - 8.4) \approx 7.4$ (comparing New England Baptist to Anna Jacques). For insurers (Figure 3.2) Blue Cross negotiates rates that are on average approximately one tenth ($\exp(-2.4 - 0) \approx 0.1$) that of Aetna and Tufts, and one third that of Fallon ($\exp(-2.4 - (-1.4)) \approx 0.37$).

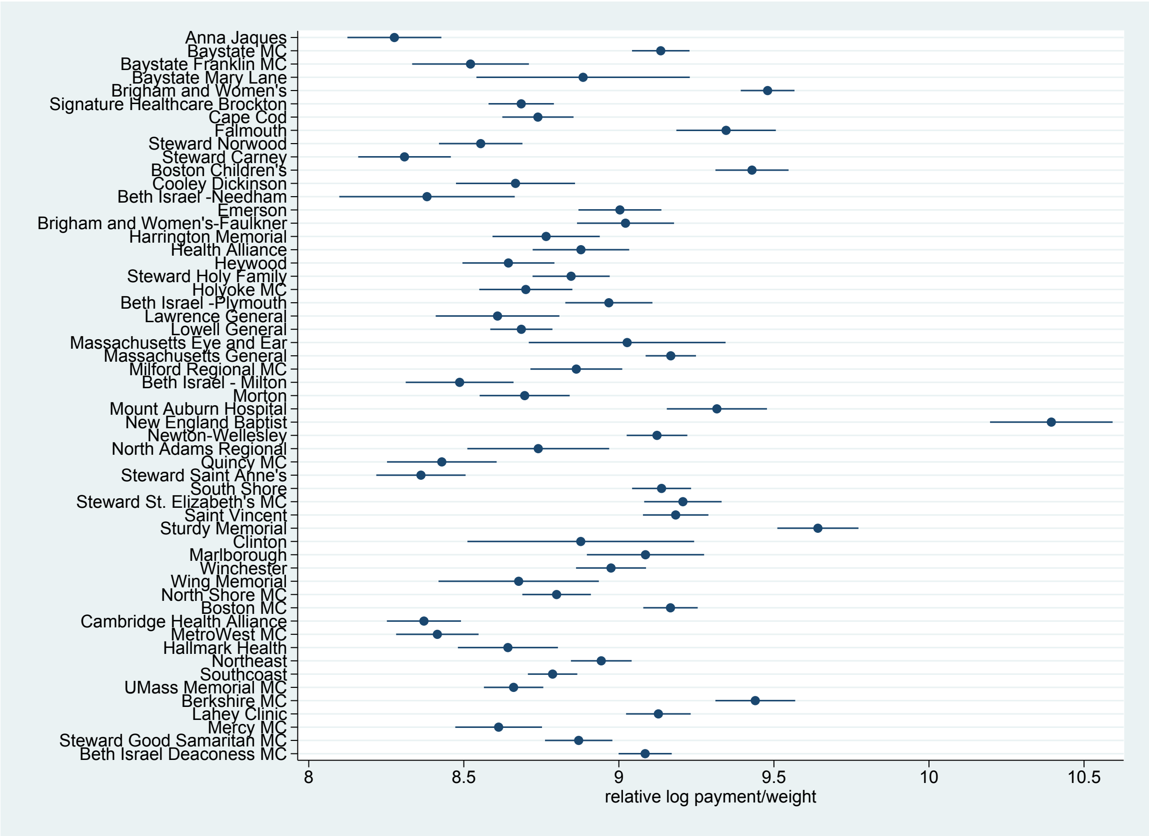


Fig. 3.1. Coefficients on hospital indicators estimated from regressing log(payment per imputed weight) against indicators for hospitals and insurers, suppressing the constant term.

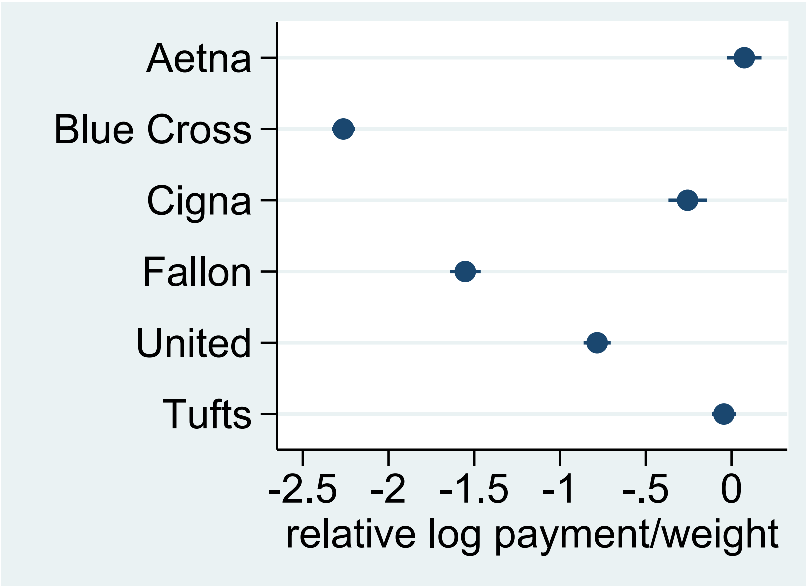


Fig. 3.2. Coefficients on insurer indicators estimated from regressing log(Payment per imputed weight) against indicators for hospitals and insurers, suppressing the constant term.

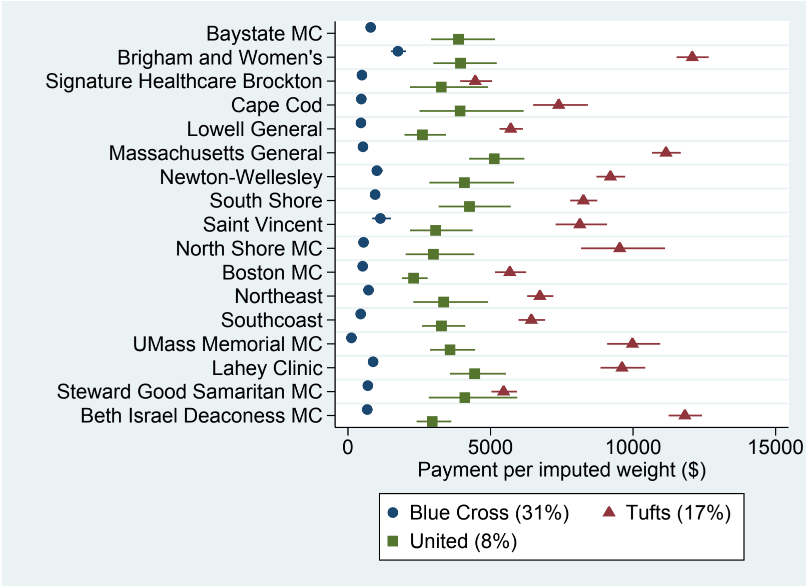


Fig. 3.3. Estimates of negotiated rates for the largest hospital-insurer pairs. Market shares are shown in parentheses (fraction of FFS admissions). The full list of negotiated rates is shown in Appendix C.3.

Appendix C.3 shows estimates of payments per imputed weight for each hospital-insurer pair with at least 30 admissions. We show these estimates for the three largest insurers in Figure 3.3. Each value in this table reports the mean payment made by an insurer for an imputed weight of 1, which we construct to be equivalent to a DRG weight of 1. Column (7) of this table highlights great variation between hospitals. For example, Berkshire Medical Center on average receives about $10,150 per imputed weight for insured individuals. This value is one of the highest in column (7), and reflects this hospital's geographic isolation: there are few substitute hospitals, so insurers are less able to direct patients living in that area elsewhere. Overall, Figure 3.3 shows that there is much more variation in compensation across insurers than across hospitals.

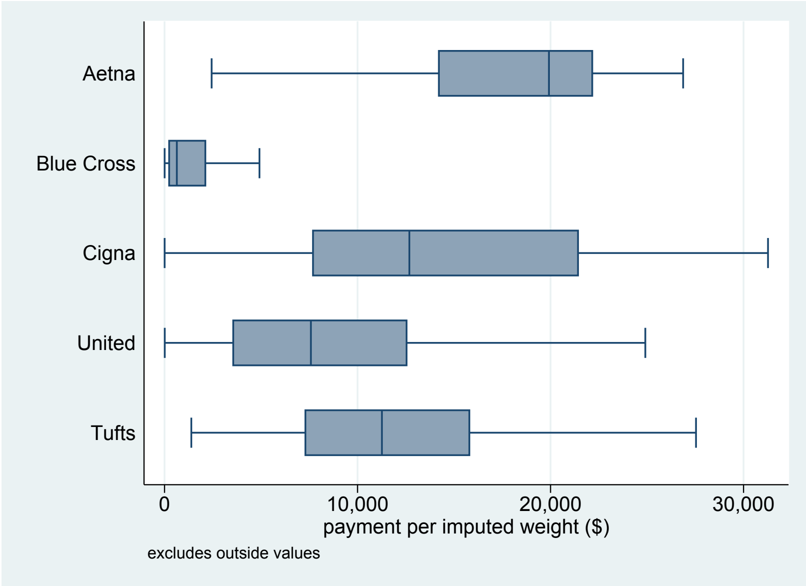


Fig. 3.4. Payment per imputed weight at Massachusetts General Hospital.

3.4.2 Within-pair variation

While previous literature has looked at hospitals as the level of observation, the patient-level data used here allows us to examine the variation within a hospital-insurer pair. In order to better understand the variation of the amount paid by both the hospital and insurance company in relation to the amount charged by the hospital, we look at the ratio of the total paid to imputed weight for each patient at a particular hospital with a particular type of insurance. Figure 3.4 shows boxplots of the ratio of payments to imputed weights for FFS payments at Massachusetts General Hospital (the largest hospital by number of FFS payments). While this figure illustrates that there is great variation in median payments for an imputed weight by insurer, there is much left to be explained. We consider the interquartile ranges shown in this figure to be economically significant, and wish to further understand this variation.

Table 3.5 investigates possible drivers of this variation by regressing the log ratio of payments per imputed weight against log negotiated rates and some patient characteristics. Column 1 shows the constant-only model, and column 2 introduces the

Table 3.5.: Regression of log payment per imputed weight against log negotiated rates estimates and patient characteristics. Column 5 restricts the dataset to only use observations that have no omitted values for the regression in Column 4.

	(1)	(2)	(3)	(4)	(5)
	log_poverwi	log_poverwi	log_poverwi	log_poverwi	log_poverwi
log(negotiated rate)		1.609***	1.593***	1.397***	1.471***
		(0.0134)	(0.0136)	(0.0242)	(0.0227)
female			0.198***	0.0331	0.0350
			(0.0187)	(0.0250)	(0.0251)
5-18 years			-0.603***	-0.575***	-0.568***
			(0.0528)	(0.0840)	(0.0842)
17-24 years			-0.428***	-0.408***	-0.366***
			(0.0497)	(0.0750)	(0.0751)
24-44 years			-0.192***	-0.249***	-0.194***
			(0.0372)	(0.0540)	(0.0539)
44-64 years			-0.409***	-0.485***	-0.439***
			(0.0364)	(0.0528)	(0.0526)
over 64 years			-0.651***	-0.671***	-0.694***
			(0.0390)	(0.0658)	(0.0521)
Other Non-Federal Programs				-1.745***	
				(0.425)	
Preferred Provider organization				0.922***	
				(0.0871)	
Point of Service				0.731***	
				(0.0856)	
Exclusive Provider Organization				0.921***	
				(0.0921)	
Indemnity Insurance				0.859***	
				(0.125)	
Health Mainienance Organization				0.712***	
				(0.0942)	
Constant	7.676***	-6.047***	-6.192***	-4.940***	-4.779***
	(0.0109)	(0.115)	(0.615)	(0.717)	(0.717)
Observations	30206	30206	30206	12396	12396
R^2	0.000	0.323	0.339	0.329	0.320
AIC	124580.6	112789.6	112144.7	41894.8	42054.3
BIC	124588.9	112806.2	112535.5	42288.3	42403.3
Race and ethnicity controls?			Y	Y	Y

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

log(negotiated rate) term. The R^2 suggests that negotiated rates account for about 32% of the variation in this ratio. Within the merged dataset we have two candidates for explaining at least some of the remaining 68%. Column 3 introduces patient demographics: age, sex, race, and ethnicity. including these characteristics improves slightly on the Column 2 regression. Of particular interest here is the positive and significant coefficient on the female indicator variable, suggesting that insurers compensate hospitals on average approximately 20% more than men for the same imputed weight. This is explored further in Column 1 of Table 3.6, which interacts sex with age categories. Here the effect goes away for all groups except for women aged between 17 and 44 years.[9] It may be that insurers negotiate differently for maternity care than for other services. However these demographic characteristics on their own account for very little variation in the paid to imputed weight ratio. In Column (2) we use only these demographics as explanatory variables, and achieve account for only 5.2% of the variation.

Another possibility is that insurers negotiate with hospitals for each of their plans, not for all of their customers together. If this was the case, then further slicing of the data by plans should provide more predictive power for payments. In the APCD we observe a variable categorizing the type of insurance plan a patient has. We use these as additional explanatory variables in Column 4 of Table 3.5. These buy marginally more explanatory power than controlling for patient demographics alone. Together, we interpret Tables 3.5 and 3.6 as suggestive evidence that a large fraction of hospitals' compensation of insured patients is explained by negotiated rates at the hospital-insurer level, but there still remains a large fraction of unexplained variation. Some of this variation may be the result of hospitals and insurers negotiating a fraction of charges for *most* services, but then negotiating other rates for some specific services, such as those provided in a maternity ward, or in a specialized care unit.

[9] Approximately 30% of women admitted in this age range are admitted for maternity care.

Table 3.6.: Regression of log payment per imputed weight against log negotiated rates estimates and patient characteristics.

	(1) log_poverwi	(2) log_poverwi
log(negotiated rate)	1.577***	
	(0.0136)	
female=1	-0.0495	-0.124
	(0.0629)	(0.0756)
5-18 years	-0.639***	-0.876***
	(0.0742)	(0.0892)
17-24 years	-0.800***	-0.977***
	(0.0770)	(0.0926)
24-44 years	-0.663***	-0.902***
	(0.0564)	(0.0678)
44-64 years	-0.450***	-0.424***
	(0.0490)	(0.0590)
over 64 years	-0.688***	-0.553***
	(0.0533)	(0.0641)
female=1 × 5-18 years	0.0840	0.0726
	(0.105)	(0.126)
female=1 × 17-24 years	0.632***	0.737***
	(0.100)	(0.121)
female=1 × 24-44 years	0.707***	1.262***
	(0.0748)	(0.0898)
female=1 × 44-64 years	0.0936	0.157
	(0.0697)	(0.0838)
female=1 × over 64 years	0.0981	0.236**
	(0.0751)	(0.0903)
Constant	-5.955***	8.343***
	(0.613)	(0.722)
Observations	30206	30206
R^2	0.344	0.052
AIC	111943.2	123074.5
BIC	112375.6	123498.6
Race and ethnicity controls?	Y	Y

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.5 Conclusion

Hospitals and insurers both have significant market power in the market for health care. Exploiting that they can influence their customers' hospital choices, insurers negotiate lower prices for services. In return, hospitals are able to boost their admissions from that insurer. The resulting negotiated rates mean that the price paid for the same services at the same hospital varies by insurer.

This paper estimates these negotiated rates, and documents the variation in negotiated rates between hospitals and insurance companies. We look at this variation within hospital-insurer pairs, with a single hospital and many insurers, with a single insurer and many hospitals and across the state of Massachusetts. We find great variation in amounts paid for services across these three dimensions. Negotiated rates appear to vary more by insurer than by hospital, suggesting that insurers' bargaining power is more heterogeneous than hospitals'. Our analysis shows that using charges as a proxy for insured patients' out-of-pocket expenses (i.e. co-insurance) may be inaccurate if the sample includes more than one insurer.

This paper builds on existing studies with the use of the Massachusetts CHIA APCD and Case Mix datasets. These allow us to analyze payments at the patient level, with rich information about both illnesses and transactions. While negotiated rates organize our data well, there remain unexplained differences in payments within a hospital-insurer pair: within any hospital-insurer pair making FFS payments, the relationship between payments and charges is not deterministic. Given our understanding of FFS payment arrangements, we conclude that negotiations happen over finer measures of service than simply a fraction of charges. Possible candidates could be (i) the most extreme case of negotiating prices for each and every service the hospital offers (e.g. minutes in the operating room, nights stayed, etc.), and (ii) negotiating differential prices over types of service, for example patients in the maternity ward could face different negotiated rates to patients in the emergency room. Controlling for patient demographics suggests that at least the latter explanation is plausible.

Further work in this area could approach this bargaining problem from a more theoretical angle. In this paper we discuss negotiation while remaining agnostic as to the bargaining game underlying the process. Our estimates of negotiated rates could be used, along with estimated models of hospital demand, to structurally estimate parameters in bargaining games. Such analysis would be useful in analyzing counterfactual claims about the market for health insurance. As both sides of the market are large, hospitals merging may increase prices, and insurers merging may lower prices. The magnitude of these changes could be predicted with such a model. Additionally, the impact of price regulation, such as those in place in Maryland, could be simulated.

REFERENCES

REFERENCES

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutstrom (2014). Discounting behavior: A reconsideration. *European Economic Review 71*, 15 – 33.

Andreoni, J. and J. Miller (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica 70*(2), 737–753.

Andreoni, J. and L. Vesterlund (2001). Which is the fair sex? gender differences in altruism. *Quarterly Journal of Economics*, 293–312.

Barberis, N. and M. Huang (2009). Preferences with frames: A new utility specification that allows for the framing of risks. *Journal of Economic Dynamics and Control 33*(8), 1555 – 1576.

Blanco, M., D. Engelmann, A. K. Koch, and H.-T. Normann (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics 13*(4), 412–438.

Bland, J. and N. Nikiforakis (2015). Coordination with third-party externalities. *European Economic Review 80*, 1–15.

Burdett, K., S. Shi, and R. Wright (2001). Pricing and matching with frictions. *Journal of Political Economy 109*(5), 1060–1085.

Cason, T. N., D. Friedman, and E. Hopkins (2010). Testing the TASP: An experimental investigation of learning in games with unstable equilibria. *Journal of Economic Theory 145*(6), 2309 – 2331.

Cason, T. N. and C. R. Plott (2014). Misconceptions and game form recognition: Challenges to theories of revealed preference and framing. *Journal of Political Economy 122*(6), pp. 1235–1270.

Cheung, Y.-W. and D. Friedman (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior 19*(1), 46 – 76.

Chmura, T., S. Kube, T. Pitz, and C. Puppe (2005). Testing (beliefs about) social preferences: Evidence from an experimental coordination game. *Economics Letters 88*, 214–220.

Chou, E., M. McConnell, R. Nagel, and C. Plott (2009). The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person guessing? game. *Experimental Economics 12*(2), 159–179.

Coller, M., G. W. Harrison, and E. E. Rutström (2012). Latent process heterogeneity in discounting behavior. *Oxford Economic Papers 64*(2), 375–391.

Conte, A., J. D. Hey, and P. G. Moffatt (2011). Mixture models of choice under risk. *Journal of Econometrics 162*(1), 79 – 88. The Economics and Econometrics of Risk.

Cooper, Z., S. V. Craig, M. Gaynor, and J. Van Reenen (2015). The price ain't right? hospital prices and health spending on the privately insured. Technical report, National Bureau of Economic Research.

Cox, J. C. and D. James (2012). Clocks and trees: Isomorphic Dutch auctions and centipede games. *Econometrica 80*(2), 883–903.

Dafny, L. (2008). Are health insurance markets competitive? Technical report, National Bureau of Economic Research.

Dafny, L., M. Duggan, and S. Ramanarayanan (2009). Paying a premium on your premium? consolidation in the us health insurance industry. Technical report, National Bureau of Economic Research.

Dafny, L., J. Gruber, and C. Ody (2015). More insurers lower premiums: Evidence from initial pricing in the health insurance marketplaces. *American Journal of Health Economics*.

Dafny, L. S. (2003). How do hospitals respond to price changes? Technical report, National Bureau of Economic Research.

Diekmann, A. (1985). Volunteer's dilemma. *The Journal of Conflict Resolution 29*(4), pp. 605–610.

Diekmann, A. (1986). Volunteers dilemma. a social trap without a dominant strategy and some empirical results. In *Paradoxical Effects of Social Behavior*, pp. 187–197. Springer.

Eeckhoudt, L., C. Gollier, and H. Schlesinger (1996). Changes in background risk and risk taking behavior. *Econometrica 64*(3), pp. 683–689.

Engelmann, D. and M. Strobel (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. *American economic review*, 857–869.

Felsö, F. A. and A. R. Soetevent (2014). Broad and narrow bracketing in gift certificate spending. *European Economic Review 66*(0), 284 – 302.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics 10*(2), 171–178.

Gaynor, M., K. Ho, and R. J. Town (2015, June). The industrial organization of health-care markets. *Journal of Economic Literature 53*(2), 235–84.

Gneezy, U. and J. Potters (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics* ., 631–645.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association 1*(1), 114–125.

Gruber, J. and R. McKnight (2016, May). Controlling health care costs through limited network insurance plans: Evidence from massachusetts state employees. *American Economic Journal: Economic Policy 8*(2), 219–50.

Harrison, G. W., M. I. Lau, and E. E. Rutström (2010). Individual discount rates and smoking: Evidence from a field experiment in denmark. *Journal of Health Economics 29*(5), 708 – 717.

Harrison, G. W., M. I. Lau, E. E. Rutström, and M. B. Sullivan (2005). Eliciting risk and time preferences using field experiments: Some methodological issues. *Field experiments in economics 10*, 125–218.

Harrison, G. W., J. A. List, and C. Towe (2007). Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion. *Econometrica 75*(2), 433–458.

Harrison, G. W. and E. Rutström, E (2009). Expected utility theory and prospect theory: one wedding and a decent funeral. *Experimental Economics 12*(2), 133–158.

Harsanyi, J. C. and R. Selten (1988, December). *A General Theory of Equilibrium Selection in Games*. Number 0262582384 in MIT Press Books. The MIT Press.

Hey, J. D. and C. Orme (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica 62*(6), 1291–1326.

Ho, K. (2006). The welfare effects of restricted hospital choice in the us medical care market. *Journal of Applied Econometrics 21*(7), 1039–1079.

Ho, K. (2008). Insurer-provider networks in the medical care market. *American Economic*.

Ho, K. and R. S. Lee (2015). Insurer competition in health care markets.

Holt, C. A. (2006). *Markets, games, and strategic behavior: recipes for interactive learning*. Pearson Addison Wesley.

Holt, C. A., S. K. Laury, et al. (2002). Risk aversion and incentive effects. *American economic review 92*(5), 1644–1655.

Julien, B., J. Kennes, and I. King (2000). Bidding for labor. *Review of Economic Dynamics 3*(4), 619 – 649.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), pp. 263–292.

Koch, A. K. and J. Nafziger (2016). Goals and bracketing under mental accounting. *Journal of Economic Theory 162*, 305 – 351.

Koop, G., D. J. Poirier, and J. L. Tobias (2007). *Bayesian econometric methods*. Cambridge University Press.

McCabe, K. A., V. L. Smith, and M. LePore (2000). Intentionality detection and mindreading?: Why does game form matter? *Proceedings of the National Academy of Sciences 97*(8), 4404–4409.

Rabin, M. and G. Weizsacker (2009). Narrow bracketing and dominated choices. *American Economic Review 99*(4), 1508–43.

Read, D., G. Loewenstein, and M. Rabin (1999). Choice bracketing. *Journal of Risk and Uncertainty 19*(1-3), 171–197.

Reinhardt, U. E. (2006). The pricing of US hospital services: chaos behind a veil of secrecy. *Health Affairs 25*(1), 57–69.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science 4*(3), 199–214.

Thaler, R. H. (2015). *Misbehaving: The making of behavioral economics.* WW Norton & Company.

Thaler, R. H., A. Tversky, D. Kahneman, and A. Schwartz (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *The Quarterly Journal of Economics 112*(2), pp. 647–661.

Tversky, A. and D. Kahneman (1981). The framing of decisions and the psychology of choice. *Science 211*(4481), 453–458.

van Dyk, D. A. and X.-L. Meng (2001). The art of data augmentation. *Journal of Computational and Graphical Statistics 10*(1), pp. 1–50.

von Gaudecker, H.-M., A. van Soest, and E. Wengstrm (2011). Heterogeneity in risky choice behavior in a broad population. *The American Economic Review 101*(2), pp. 664–694.

Wilcox, N. T. (2006). Theories of learning in games and heterogeneity bias. *Econometrica 74*(5), pp. 1271–1292.

APPENDICES

A. APPENDIX FOR: HOW MANY GAMES ARE WE PLAYING? AN EXPERIMENTAL ANALYSIS OF CHOICE BRACKETING IN GAMES

A.1 Equilibrium analysis

Consider the two games in Table A.1. Games Γ_1 and Γ_2, in isolation, are both 2-player Volunteer's Dilemmas (Diekmann, 1985, 1986). In Γ_1, choosing action A guarantees earnings of a, while choosing B yields earnings of x iff the other player does not choose A. Similar incentives exist in Γ_2. As earnings are additive across the two tables, a (mixed) strategy is a probability distribution over all four possible combinations of actions $\{A, B\} \times \{C, D\}$, and assuming broad bracketing, standard game theory would treat this game no differently to the 4-action, 2-player game presented in Table A.2.

For the remainder of this paper, I use p_{ij} to denote the joint probability of taking action i in Γ_1 and j in Γ_2, and p_k to denote the unconditional probability of taking

Table A.1.: The Roommate's Dilemma. Players chose a pair of actions $\{AC, AD, BC, BD\}$, payoffs are the sum of outcomes in both tables. $x > a > 0$, $y > c > 0$, $a < c$.

Γ_1

	A	B
A	a, a	a, x
B	x, a	$0, 0$

Γ_2

	C	D
C	c, c	c, y
D	y, c	$0, 0$

Table A.2.: Broad-bracketed version of Γ_1.

	AC	AD	BC	BD
AC	$a+c$, $a+c$	$a+c$, $a+y$	$a+c$, $x+c$	$a+c$, $x+y$
AD	$a+y$, $a+c$	a, a	$a+y$, $x+c$	a, x
BC	$x+c$, $a+c$	$x+c$, $a+y$	c, c	c, y
BD	$x+y$, $a+c$	x, a	y, c	0, 0

Table A.3.: Characterization of equilibria based on unpaired choice probabilities in bracketed games.

		Player 1		Player 2	
		p_A	p_C	p_A	p_C
Pure actions only	1	0	0	1	1
	2	0	1	1	0
	3	1	0	0	1
	4	1	1	0	0
One pure, one mixed	5	a/x	0	a/x	1
	6	a/x	1	a/x	0
	7	0	c/y	1	c/y
	8	1	c/y	0	c/y
Both mixed	9	a/x	c/y	a/x	c/y

action k in its relevant game (i.e. $p_A = p_{AC} + p_{AD}$). When discussing asymmetric equilibria, I use superscripts to denote player labels.

A.1.1 Risk-neutral equilibria of the broad-bracketed game

Note that there are 3 Nash equilibria in each bracketed game, and therefore in the broad-bracketed game there are $3 \times 3 = 9$ types of equilibria, based on whether players mix in zero, one, or both bracketed games. These are summarized in Table A.3. Note that apart from the equilibria where players mix in both games, these uniquely categorize the equilibrium. For example, for the equilibrium described in row 5, players mix in game 1 and don't mix in game 2, so it must be that:

$$p^1_{AD} = a/x,\ p^1_{BD} = 1 - a/x$$
$$p^2_{AC} = a/x,\ p^2_{BC} = 1 - a/x$$

and all other actions are played with zero probability. However, for the final case (case 9 in Table A.3) where both players mix in both games, a continuum of equilibria exist.

First I characterize the set of strategies that make players indifferent between taking all four action pairs. Note that since playing AC always yields a payoff of

$a + c$, any fully mixed strategy equilibrium must achieve this payoff. Therefore we can characterize any such MSNE as a pair[1] of $\mathbf{p}^*$ satisfying:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ a+c & a+c & a+c & a+c \\ a+y & a & a+y & a \\ c+x & c+x & c & c \\ x+y & x & y & 0 \end{bmatrix} \mathbf{p}^* = \begin{bmatrix} 1 \\ a+c \\ a+c \\ a+c \\ a+c \end{bmatrix} \tag{A.1}$$

which has solutions:

$$p_{AC} - p_{BD} = \frac{a\,y + c\,x - x\,y}{x\,y} \tag{A.2}$$

$$p_{AD} + p_{BD} = 1 - c/y \tag{A.3}$$

$$p_{BC} + p_{BD} = 1 - a/x \tag{A.4}$$

Note that (A.3) and (A.4) are the MSNE probabilities of each bracketed game. Expressing these in terms of the SOSD choice p_{AD}:

$$\mathbf{p}^* = \begin{bmatrix} p_{AC} \\ p_{AD} \\ p_{BC} \\ p_{BD} \end{bmatrix} = \begin{bmatrix} a/x + c/y - 1 + 1 - c/y - p_{AD} \\ p_{AD} \\ 1 - a/x - (1 - c/y - p_{AD}) \\ 1 - c/y - p_{AD} \end{bmatrix} = \begin{bmatrix} a/x - p_{AD} \\ p_{AD} \\ c/y - a/x + p_{AD} \\ 1 - c/y - p_{AD} \end{bmatrix} \tag{A.5}$$

Next I focus on strategies that exclude exactly one action pair. These strategies must be nested within (A.5). Specifically, the strategy must still make his opponent's expected payoff from the bracketed games equal to a and c respectively, and hence $a+c$ in the full game. Otherwise she has a pure-strategy best response, and therefore the column player must not be willing to mix. Therefore any mixed strategy Nash

[1] As there are many solutions to (A.1), all of which make opponents indifferent between the four action pairs, it suffices that strategies are played that satisfy this equation, and it is not necessary that both players use the same strategy.

equilibrium strategy that puts positive probability on exactly three action pairs must also satisfy (A.5).

If an equilibrium strategy puts positive probability on only two action pairs, then it must be that these action pairs correspond to a pure strategy in one bracketed game, and the mixed strategy in the other. First, note that the equilibria described above are the same as types 5-8 in Table A.3. To see that we cannot have other equilibria of this type, suppose that the column player sets $p_{AD} = p_{BC} = 0$, so only ever plays actions AC and BD. This fixes the probability that column plays A to the same value as the probability she plays C, which means that row's best response involves playing a pure strategy in at least one of the bracketed games. If column best responded to this, she would not be willing to play here candidate strategy. This argument also rules out strategies setting $p_{BC} = p_{AD} = 0$.

A.1.2 Some possibly salient equilibria

Payoff and risk dominance Harsanyi and Selten (1988) identify payoff and risk dominance as arguments for selecting among multiple equilibria. As the Roommate's Dilemma is an anti-coordination game, however, these do dot have much bite. Payoff dominance suggests that a pure-strategy Nash equilibrium will be selected. With random matching in the experiment design, however, it is unlikely that subjects will successfully coordinate on pure strategies. Risk dominance suggests favor for actions A and C if $\frac{a}{x} > \frac{1}{2}$ and $\frac{c}{y} > \frac{1}{2}$ respectively, but players could not all take these actions in equilibrium.

Symmetry Another appealing selection criterion is symmetry, which suggests that both players' choice probabilities are the same.[2] This rules out all pure-strategy Nash equilibria and all mixed strategy Nash equilibria except for those where both players

[2]This selection criterion has been used extensively in directed search games, which include an anti-coordination game in the second stage. See for example Julien et al. (2000); Burdett et al. (2001)

choose according to (A.5), parameterized by the same p_{AD}. Symmetry has appeal in this setting due to the random matching in the experiment design.

Independence One could further restrict the symmetric equilibria to the equilibrium where actions are independent across the bracketed games. The choice probabilities would therefore be:

$$\mathbf{p}^* = \begin{bmatrix} \frac{a}{x}\frac{c}{y} \\ \frac{a}{x}\left(1-\frac{c}{y}\right) \\ \left(1-\frac{a}{x}\right)\frac{c}{y} \\ \left(1-\frac{a}{x}\right)\left(1-\frac{c}{y}\right) \end{bmatrix}$$

which is a special case of (A.5). Note that this equilibrium places positive weight on the SOSD action AD.

Exclusion of SOSD actions If players are broad bracketers and have only a very small degree of risk aversion, then they may play strategies close to:

$$\mathbf{p}^* = \begin{bmatrix} \frac{a}{x} \\ 0 \\ \frac{c}{y}-\frac{a}{x} \\ 1-\frac{c}{y} \end{bmatrix}$$

Inequality Chmura et al. (2005) and **?** demonstrate that inequality can matter in coordination games, and therefore equilibria may be selected that result in a more equitable distribution of earnings. This would suggest that the (AD, BC) and (BC, AD) equilibria (i.e. where each player takes one "safe" action and one "risky" action) may be more likely to be selected over equilibria where one player takes both safe actions, and hence earns much less than their opponent.

A.1.3 Narrow bracketing, second-order stochastic dominance, and background risk

Until this point, I have restricted attention to predictions made by assuming risk neutrality. I now change focus and investigate the implications of risk averse preferences in the game. Specifically, I identify the action pairs that can be ruled out by second-order stochastic dominance:

Definition A.1.1 *Distributions $F(x)$* **second-order stochastically dominates** *(SOSD) $G(x)$ iff:*

$$\int_{-\infty}^{x} F(w) - G(w)dw \leq 0 \ \forall x \in \mathbb{R}$$

and there exists an $x \in \mathbb{R}$ such that the above inequality is strict.

Given the choice between lotteries, no risk-averse individual would choose one that was SOSD by another.

Figure A.1 shows the cumulative density functions of the lotteries induced by the four action pairs. Using the restrictions:

$$a < c, \quad a < x, \quad c < y \tag{A.6}$$

which will be maintained in all treatments, then one can derive the conditions on subjective beliefs (p_A, p_B) such that one SOSD another by checking that the difference in the integrals up to $x + y$ is at most zero.

Theorem A.1.1 *Given subjective beliefs (p_A, p_B) :*

1. *AX SOSD BX $\forall X \in \{C, D\}$ iff $p_A \leq \frac{a}{x}$*
2. *XC SOSD XD $\forall X \in \{A, B\}$ iff $p_C \leq \frac{c}{y}$*

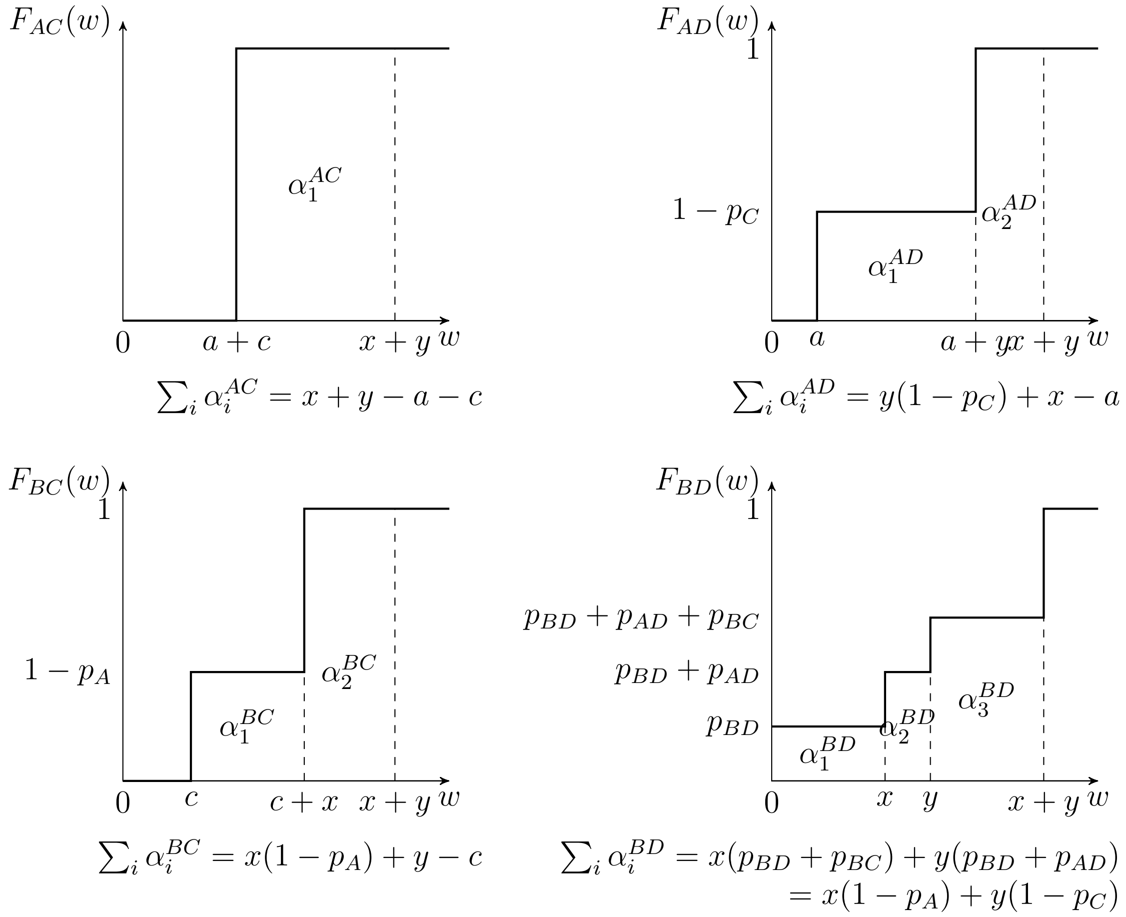


Fig. A.1. Cumulative density function for the four action pairs (not to scale). $\int_{-\infty}^{x+y} F(w)dw$ is shown below each plot, which is used to calculate conditions for second-order stochastic dominance.

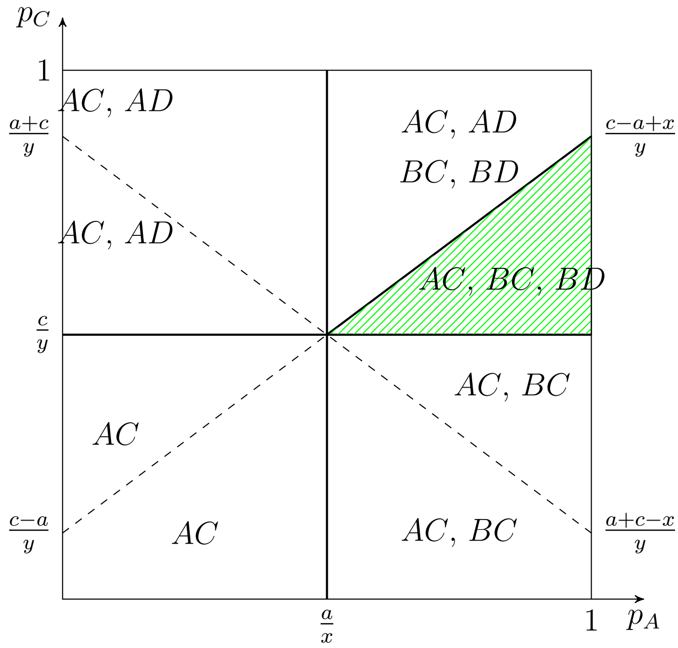


Fig. A.2. Actions permissible by risk-averse individuals. In the shaded region, a narrow-bracketing individual may choose AD mistakingly. All boundaries intersect at the risk-neutral mixed-strategy Nash equilibrium. Shown for parameterizations satisfying $a + x < c + y$.

3. *BC SOSD AD iff*

$$p_C \leq \begin{cases} \frac{c-a+p_A x}{y} & \text{if } c+x \leq a+y \\ \frac{c-a+p_A(a-c+y)}{y} & \text{otherwise} \end{cases}$$

4. *AC SOSD BD iff* $p_C \leq \frac{a+c-p_A x}{y}$

Note that inequalities 1, 2, and 4 bind at the risk-neutral mixed-strategy Nash equilibrium $(p_A^*, p_B^*) = (a/x, c/y)$. Therefore if risk-averse players were playing against a risk-neutral Nash equilibrium strategy, AC is their unique best response. Inequality 3 binds in the mixed-strategy Nash equilibrium iff $a + x \leq c + y$.

While not directly expressed in the payoff tables, transferring earnings between Γ_1 and Γ_2, or simply adding or subtracting earnings from one table affects play differently depending on the assumptions made about bracketing.

Theorem A.1.2 *Hold a player's beliefs about opponents' strategies constant. If a subject has decreasing absolute risk aversion, then:*

1. *Subtracting $z > 0$ from all payoffs in Γ_1 and adding z to all payoffs in Γ_2 (i) will not affect the player's decisions if she broadly brackets ($\kappa = 1$), and (ii) will make a narrow bracketer ($\kappa < 1$) more likely to take action A holding their C/D choice constant, and more likely to take action D (holding their A/B choice constant).*
2. *Subtracting $z > 0$ from Γ_1 (i) will make the player more likely to choose A (holding their C/D choice constant) irrespective of their bracketing behavior, (ii) iff the player is not a narrow bracketer ($\kappa = 0$) will make them more likely to choose C (holding their A/B) choice constant.*

These results follow directly from Eeckhoudt et al. (1996, See their Proposition 1, p685). Transferring earnings from one table to another does not affect broad bracketers because they always "see" the payoffs as in Table A.2. Changing the earnings of one game while leaving the other unchanged does not change narrow bracketers' behavior in the unchanged game because they ignore these changes when making their decision in this game.

A.1.4 Risk aversepredictions

Here I state some results, derived in Appendix A.2, that may be used to distinguish narrow and broad bracketing play in equilibrium in the Roommate's Dilemma, and for behavior in the Lottery Task:

Theorem A.1.3 *The utility (and hence certainty equivalent) of any non-degenerate gamble is strictly decreasing in bracketing parameter κ (see Equation 1.1).*

Theorem A.1.4 *Any risk-averse player who does not fully narrow bracket (i.e. $\kappa > 0$ in (1.1)) cannot be made indifferent between taking all four actions.*

It follows from this that the only way a risk-averse player can appear to be indifferent between all four actions is if she narrowly brackets ($\kappa = 0$).

Theorem A.1.5 *Hold DARA risk preferences constant. The tokens invested in Part 2 of the Lottery Task is increasing in κ (see Equation 1.1).*

A.2 Proofs

Proof of Theorem A.1.1

Proof 1. Since gamble X is a common element of both AX and BX, it is sufficient to show that A SOSD B for $p_A \leq \frac{a}{x}$. A SOSD B requires:

$$\begin{aligned} 0 &\geq \int_{-\infty}^{x+y} F_A(w) - F_B(w)dw \\ &= (x+y-a) - [x(1-p_A) + y] \\ &= xp_A - a \\ \iff p_A &\leq \frac{a}{x} \end{aligned}$$

2. Since gamble X is a common element of both XC and XD, it is sufficient to show that C SOSD D for $p_C \leq \frac{c}{y}$. C SOSD D requires:

$$\begin{aligned} 0 &\geq \int_{-\infty}^{x+y} F_C(w) - F_D(w)dw \\ &= (x+y-c) - [y(1-p_C) + x] \\ &= yp_C - c \\ \iff p_C &\leq \frac{c}{y} \end{aligned}$$

3. For the case where $c + x \leq a + y$, and using the relevant areas in Figure A.1, we require:

$$
\begin{aligned}
0 &\geq \int_{-\infty}^{x+y} F_{BC}(w) - F_{AD}(w)dw \\
&= x(1 - p_A) + y - c - [y(1 - p_C) + x - a] \\
&= a - c - xp_A + yp_C \\
\iff p_C &\leq \frac{c - a + xp_A}{y}
\end{aligned}
$$

for the case where $c + x > a + y$, the cdfs cross twice. First at c, and again at $a + y$. We therefore need to check that:

$$
\begin{aligned}
0 &\geq \int_{-\infty}^{x+y} F_{BC}(w) - F_{AD}(w)dw \\
&= -(c - a)(1 - p_C) + (p_C - p_A)(a + y - c) \\
&= -c + a - p_A(a - c + y) + p_C y \\
\iff p_C &\leq \frac{c - a + p_A(a - c + y)}{y}
\end{aligned}
$$

4. Using the relevant areas in Figure A.1, we require:

$$
\begin{aligned}
0 &\geq \int_{-\infty}^{x+y} F_{AC}(w) - F_{BD}(w)dw \\
&= x + y - a - c - [x(1 - p_A) + y(1 - p_C)] \\
&= xp_A + yp_C - a - c \\
\iff p_C &\leq \frac{a + c - xp_A}{y}
\end{aligned}
$$

■

Proof of Theorem A.1.3

Proof Let Z be a multivariate random variable with distribution function $F(z)$ and marginals $L_i(z_i)$. The monetary payoff from realization z is $\sum_i z_i$. If an agent with utility specified in (1.1), and brackets according to each element of X, then they have an expected payoff for this lottery of:

$$
\begin{aligned}
U(\kappa) &= \kappa \int u\left(\sum_i z_i\right) dF(x) + (1-\kappa) \sum_i \int u(z_i) dL_i(z_i) \\
U'(\kappa) &= \int u\left(\sum_i z_i\right) dF(z) - \sum_i \int u(z_i) dL_i(z_i) \\
&< \int \sum_i u(z_i) dF(z) - \sum_i \int u(z_i) dL_i(z_i) = 0
\end{aligned}
$$

where the inequality follows by concavity of u. As utility is decreasing in κ, it must be that the certainty equivalent is decreasing in κ for all non-degenerate, bracketed gambles. ■

Proof of Theorem A.1.4

Proof Suppose to the contrary that the player is made indifferent between taking all four actions. Because the player is risk-averse and AC yields certain earnings of $a+c$, it follows that actions AD, BC, and BD must yield a higher monetary payoff in expectation. Suppose therefore that the expected monetary payoffs are:

$$
E = \begin{bmatrix} a+c \\ a+c+\delta_1 \\ a+c+\delta_2 \\ a+c+\delta_3 \end{bmatrix}
$$

with all δs strictly positive. The risk-averse MSNE probabilities $\mathbf{p}^r$ must therefore be characterized by the system of equations:

$$
\left[\begin{array}{cccc|c}
a+c & a+c & a+c & a+c & a+c \\
a+y & a & a+y & a & a+c+\delta_1 \\
c+x & c+x & c & c & a+c+\delta_2 \\
x+y & x & y & 0 & a+c+\delta_3
\end{array}\right]
$$

$$
\sim \left[\begin{array}{cccc|c}
1 & 1 & 1 & 1 & 1 \\
y-c & -c & y-c & -c & \delta_1 \\
x-a & x-a & -a & -a & \delta_2 \\
x+y-a-c & x-a-c & y-a-c & -a-c & \delta_3
\end{array}\right]
\qquad
\begin{array}{l}
\div(a+c) \\
-[1] \\
-[1] \\
-[1]
\end{array}
$$

$$
\sim \left[\begin{array}{cccc|c}
1 & 1 & 1 & 1 & 1 \\
0 & -y & 0 & -y & \delta_1-y+c \\
0 & 0 & -x & -x & \delta_2-x+a \\
0 & -y & -x & -x-y & \delta_3-x-y+a+c
\end{array}\right]
\qquad
\begin{array}{l}
\\
-(y-c)\times[1] \\
-(x-a)\times[1] \\
-(y+x-a-c)\times[1]
\end{array}
$$

$$
\sim \left[\begin{array}{cccc|c}
1 & 1 & 1 & 1 & 1 \\
0 & -y & 0 & -y & \delta_1-y+c \\
0 & 0 & -x & -x & \delta_2-x+a \\
0 & 0 & 0 & 0 & \delta_3
\end{array}\right]
\qquad
\begin{array}{l}
\\
\\
\\
-[2]-[3]
\end{array}
$$

$$
\sim \left[\begin{array}{cccc|c}
1 & 1 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{array}\right]
\qquad
\begin{array}{l}
-[4]\div\delta_3 \\
([2]-[4]\times(\delta_1-y+c)/\delta_3)\div(-x) \\
([3]-[4]\times(\delta_2-x+a)/\delta_3)\div(-x) \\
\div\delta_3
\end{array}
$$

$$
\sim \left[\begin{array}{cccc|c}
1 & 0 & 0 & -1 & 0 \\
0 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{array}\right]
\qquad
\begin{array}{l}
-[2]-[3] \\
\\
\\
\\
\end{array}
$$

Note that the last row of this matrix is equivalent to $\mathbf{0}'\mathbf{p} = 1$, but $\mathbf{p} \in \Delta^3$, a contradiction,[3] and so at most one action must be played with zero probability in equilibrium.

■

[3]This does not work for the risk-neutral case because all of the δs are equal to zero and we must divide by δ_3

Proof of Theorem A.1.5

Proof Fix the amount for the blue investment at x_B and consider the restricted problem of choosing only x_R with x_B fixed. A broad bracketer notices that the B investment is background risk for the problem of choosing x_R, and that this background risk is a sure payment of $1 - x_B$, plus 50% chance of winning χx_B, which first-order stochastically dominates receiving nothing for sure, and independent of the lottery induced by her choice of x_R. Therefore by Eeckhoudt et al. (1996, p685, Theorem 1), the broad bracketing subject behaves uniformly less risk-averse by noticing this background risk that a narrow bracketer would ignore. The broad bracketer's choice of x_R conditional on x_B is hence higher than that of the narrow bracketer.

The above reasoning has established that $x^*(\kappa = 0) < x^*(\kappa = 1)$. It remains to show that x^* is monotonic is κ. To see this, note that we can write a κ-bracketer's utility as:

$$U_\kappa(x) = \kappa U^B(x) + (1 - \kappa)U^N(x)$$

where the B and N superscripts denote the narrow and broad bracketing extremes. As U^B and U^N are globally concave, it follows that $x^*(\kappa) \in [x^*(0), x^*(1)]$, and by induction for $\kappa' < \kappa''$ that $x^*(\kappa') \in [x^*(0), x^*(\kappa'')]$. x^* must therefore be increasing in κ. ■

A.3 Notes on Bayesian estimations

A.3.1 Prior distributions

Population parameters $(\beta, \Sigma, \lambda^{RD}, \lambda^{LT})$ require specification of priors. These are discussed below.

I use a normal-inverse Wishart prior β and Σ. For β, I set the mean of the first row (i.e. the constant term) equal to reasonable expected values for the relevant parameters, and any remaining rows (if there are covariates) equal to zero:

$$\underline{\mu}_\beta = \begin{bmatrix} \log(0.5) - 0.5 & \log(0.5) - 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ & \vdots & & \vdots & \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{A.7}$$

and set the prior variance of each element of β to 1. This achieves prior means of 0.5 for the risk aversion parameters, and uniform priors on γ_i and the probability of bracketing in each task.

For the upper-left block of Σ, which I denote $\Sigma_{-\kappa}$ (covariance of $(r_i^{RD}, r_i^{LT}, \gamma_i)$), I assume:

$$\Sigma_{-\kappa}^{-1} \sim \text{Wishart}(5, 5\mathcal{I}_3) \tag{A.8}$$

Where $\mathcal{I}$ is the identity matrix.

For the lower-right block of Σ, which I denote Σ_κ, there is one unrestricted parameter, ω, which is the covariance of the two latent type processes. I assume:

$$\omega \sim \mathcal{U}[-1, 1] \tag{A.9}$$

Finally, I assume log-normal priors for the logit precision parameters:

$$\log(\lambda^{RD}) \sim \mathcal{N}(-3, 1) \tag{A.10}$$

$$\log(\lambda^{LT}) \sim \mathcal{N}(-4, 2) \tag{A.11}$$

A.3.2 Likelihood function

This section derives the likelihood function and describes the data augmentation process used to simulate the likelihood.[4]

Let Y denote the data obtained from the experiment. This contains all choices made by all subjects in both the Roommate's Dilemma, and the Lottery Task. Subject i's action in period t of the Roommate's Dilemma is $a_{i,t}$, and their decision in instance s of the lottery task is $b_{i,s}$. Y can therefore be defined as:

$$Y \equiv \bigcup_i \left\{ \{a_{i,t}\}_{t=1}^{20} \cup \{b_{i,s}\}_{s=1}^{8} \right\} \tag{A.12}$$

I augment these data by the individual-specific parameters $Z_i = (r_i^{RD}, r_i^{LT}, \kappa_i^{RD}, \kappa_i^{LT}, \gamma_i)$ to achieve:

$$Y^* \equiv \bigcup_i \left\{ \{a_{i,t}\}_{t=1}^{20} \cup \{b_{i,s}\}_{s=1}^{8} \cup \{r_i^{RD}, r_i^{LT}, \kappa_i^{RD}, \kappa_i^{LT}, \gamma_i\} \right\} \tag{A.13}$$

In the remainder of this section, I first derive the likelihood functions conditional on the augmented data Y^* and the model parameters θ, where:

$$\theta \equiv (\lambda_{RD}, \lambda_{LT}, \mu_r^{RD}, \nu_r^{RD}, \mu_r^{LT}, \nu_r^{LT}, \rho^{RD}, \rho^{LT}, \mu_\gamma, \nu_\gamma) \tag{A.14}$$

I then derive the joint posterior distribution of the parameters and latent variables. Finally, I outline the procedure for drawing from this distribution.

Roommate's Dilemma likelihood function: With the logistic choice rule, if action a yields utility $U(a)$, then it will be played with probability proportional to $\exp(\lambda_{RD} U(a))$. As parameters $(r_i^{RD}, \kappa_i^{RD}, \gamma_i)$ and opponents' previous actions $\{a_{j,\tau}\}_{j \in G_i, \tau < t}$ (G contains the list of subjects in i's group) completely determine ex-

[4] For an outline of data augmentation, see van Dyk and Meng (2001).

pected utility, we can write the likelihood of observing sequence $\{a_{i,t}\}_{t=1}^{20}$, conditional on the individual-specific parameters, in the Roommate's Dilemma as:

$$p_i^{RD}(Y_i^* \mid \theta) = p_i^{RD}(Y_i \mid \theta, Z_i) p(Z_i \mid \theta) \tag{A.15}$$

$$= \prod_t \left[\frac{\exp\left[\lambda_{RD} U(a_{i,t}; \{a_{j,\tau}\}_{j \in G_i, \tau < t}, r_i^{RD}, \kappa_i^{RD}, \gamma_i)\right]}{\sum_a \exp\left[\lambda_{RD} U(a; \{a_{j,\tau}\}_{j \in G_i, \tau < t}, r_i^{RD}, \kappa_i^{RD}, \gamma_i)\right]} \right] p(Z_i \mid \theta) \tag{A.16}$$

where expected utilities $U(\cdot)$ are calculated using CRRA parameter r_i^{RD}, bracketing parameter κ_i^{RD} and beliefs parameter γ_i. Note that, conditional on the individual-specific parameters, none of the parameters governing the distribution of these enter the above equation.

Lottery Task likelihood function: as with choices in the Roommate's Dilemma, the probability of observing choices $\{b_{i,s}\}$ in the Lottery Task, conditional on all parameters including the individual-specific parameters, is:

$$p_i^{LT}(Y_i^* \mid \theta) = p_i^{LT}(Y_i \mid \theta, Z_i) p(Z_i \mid \theta) \tag{A.17}$$

$$= \prod_s \left[\frac{\exp\left[\lambda_{LT} V_s(b_{i,s}; r_i^{LT}, \kappa_i^{LT})\right]}{\sum_b \exp\left[\lambda_{LT} V_s(b; r_i^{LT}, \kappa_i^{LT})\right]} \right] p(Z_i \mid \theta) \tag{A.18}$$

where $V_s(b; \cdot)$ is the expected utility of taking action b in instance s of the lottery task. As the choice set is large for the lottery task (101 and 101^2 for parts 1 and 2 respectively), I use two approximations of this function: Firstly, in the second part where subjects make a decision for both the red lottery and the blue lottery, I take the mean of the two decisions and calculate the expected utility as if they made this mean choice in both lotteries. Secondly, I linearly interpolate (A.18) over 21 evenly spaced grid points so that the numerator need only be evaluated at multiples of 5. Note that at any two grid points, the probability ratio property of logit choice is preserved.

Simulation of the posterior distribution

Let $L_i^* = (r_i^{RD}, r_i^{LT}, \kappa_i^{*,RD}, \kappa_i^{*,LT}, \gamma_i)$ be the latent, subject-specific variables in the model, and $Y^* = Y \cup L^*$. I use the following result to derive the joint posterior distribution of (θ, Z), up to a normalizing constant:

$$p(\theta, Z \mid Y) \propto p(Y, Z \mid \theta)p(\theta) = p(Y^* \mid \theta)p(\theta) \tag{A.19}$$

$$= \prod_{\forall i} \left[p_i^{RD}(Y_i^* \mid \theta)p_i^{LT}(Y_i^* \mid \theta)p(Z_i \mid \theta)\right] p(\lambda_{RD})p(\lambda_{LT})p(\beta, \Sigma) \tag{A.20}$$

where (A.19) applies Bayes' theorem, and (A.20) substitutes in expressions of the likelihood ($p(Y^* \mid \theta)$, given by (A.16) and (A.18)) and the prior ($p(\theta)$). Note that the term $p(Z_i \mid \theta)$ appears only once in (A.20), but it appears both in (A.16) and (A.18). This is because the variable Z_i is common to both the Roommate's Dilemma and the Lottery Task. The product term in (A.20) follows because of the maintained assumption of independence between subjects. I leave (A.20) with all hyperparameters governing the distribution of individual parameters lumped into the same prior $p(\beta, \Sigma)$. While some estimations assume no dependence between an individual's parameters, I derive the general case here where there may be dependence.

Conditional distributions of latent parameters $(r_i^{RD}, r_i^{LT}, \gamma_i)$**:** All latent variables take on non-standard posterior distributions. I therefore use the Metropolis-Hastings algorithm to draw each parameter. This algorithm requires only that the distribution function is known up to a constant of proportionality. Note from (A.20) that this is exactly what we have.

For the step to simulate $z_i \in \{r_i^{RD}, r_i^{LT}, \gamma_i\}$, we require the distribution, up to proportionality, of:

$$z_i \mid Y, Z_{-z_i}, \theta$$

where "Z_{-z_i}" denotes all elements of Z except for z_i. Inspection of (A.20) reveals that no element of θ appears in this, and so $p(z_i \mid Y, (Z - z_i))$ is proportional to the element of the likelihood function specific to individual i, that is:

$$p(z_i \mid Y, (Z - z_i), \theta) \propto p_i^{RD}(Y_i^* \mid \theta) p_i^{LT}(Y_i^* \mid \theta) p(Z_i \mid \theta) \tag{A.21}$$

For each z_i, a proposal z_i' is drawn from density $q(z_i' \mid z_i)$ satisfying $q(z_i' \mid z_i) = q(z_i \mid z_i')$, z_i is then updated as follows:

$$z_i^{\text{next}} = \begin{cases} z_i' & \text{if } \frac{p(z_i' \mid Y, (Z - z_i), \theta)}{p(z_i \mid Y, (Z - z_i), \theta)} > u \\ 0 & \text{otherwise} \end{cases} \tag{A.22}$$

where $u \sim \mathcal{U}[0, 1]$.[5]

Conditional distributions of latent parameters ($\kappa_i^{RD}, \kappa_i^{LT}$)**:** Unlike the other latent parameters, the conditional distributions of the bracketing parameters can be derived. (A.21) still applies, but further simplification can be gained by noting that if ρ, the fraction of broad bracketers is known, the distribution of κ can be derived exactly:

$$\kappa_i^W \sim \text{Bernoulli}\left(\frac{\rho p_i^W(Y_i^* \mid \theta)\big|_{\kappa_i = 1}}{\rho p_i^W(Y_i^* \mid \theta)\big|_{\kappa_i = 1} + (1 - \rho) p_i^W(Y_i^* \mid \theta)\big|_{\kappa_i = 0}} \right) \tag{A.23}$$
$$\text{for } W \in \{RD, LT\}$$

Conditional distribution of logit choice precision parameters ($\lambda_{RD}, \lambda_{LT}$)**:** Noting that each λ only appears in the likelihood specific to its task, and its prior,

[5]For all variables I use a normal density with mean z_i and variance tailored to each parameter. In practice, I determine the inequality in (A.22) using logs because the likelihood functions are all calculated on a log scale, so it is computationally more convenient to log u rather than exponentiate the likelihoods.

the conditional posterior distribution for each logit precision parameter is proportional to:

$$p(\lambda_W \mid Y^*, (\theta - \lambda_W)) \propto p(\lambda_W) \prod_{\forall i} p_i^W(Y^* \mid \theta), \quad \text{for } W \in \{RD, LT\} \tag{A.24}$$

As this does not conform to a known distribution, I use Metropolis-Hastings steps to draw each λ.

Conditional distributions of $(\mu_r, \nu_r, \mu_\gamma, \nu_\gamma)$: Both r_i (from each task) and $\Phi^{-1}(\gamma_i)$ are assumed to be independent draws from normal distributions. Inspection of (A.20) shows that conditional on the individual parameters, the posteriors of these hyper-parameters are conjugate. That is:

$$p(\mu_r, \nu_r \mid Y^*, \theta_{-(\mu_r,\nu_r)}) \propto \prod_i p(Z_i \mid \theta) p(\beta, \Sigma) \tag{A.25}$$

It follows that:

$$\mu_r \mid Y^*, \theta_{-\mu_r} \sim \mathcal{N}\left[\frac{\sum_i r_i/\nu_r + \underline{E}(\mu_r)/\underline{V}(\mu_r)}{(N/\nu_r + \underline{V}(\mu_r)^{-1})^{-1}}, (N/\nu_r + \underline{V}(\mu_r)^{-1})^{-1}\right] \tag{A.26}$$

$$\nu_r \mid Y^*, \theta_{-\mu_r} \sim \mathcal{IG}\left(\underline{a} + N/2, \left(\underline{b}^{-1} + \frac{1}{2}\sum_i (r_i - \mu_r)^2\right)^{-1}\right) \tag{A.27}$$

where underlines denote prior parameters. The same applies for (μ_γ, ν_γ).

Conditional distribution of ρ: Inspection of (A.20) shows that conditional on the individual parameters, the posterior density of ρ (for each task W) is proportional to:

$$p(\rho^W \mid \theta_{-\rho^W}, Y^*) \propto \prod_i p(Z_i \mid \theta) p(\beta, \Sigma) \tag{A.28}$$

$$\propto \prod_i (\rho^W)^{\kappa_i^W} (1 - \rho^W)^{1-\kappa_i^W} I(\rho^W \in [0,1]) \tag{A.29}$$

$$= (\rho^W)^{\sum_i \kappa_i^W} (1 - \rho^W)^{\sum_i (1-\kappa_i^W} I(\rho^W \in [0,1]) \tag{A.30}$$

where (A.29) substitutes in the uniform prior on ρ^W. (A.30) is the Beta kernel, therefore:

$$\rho^W \mid \theta_{-\rho^W}, Y^* \sim \text{Beta}\left(\sum_i \kappa_i^W, \sum_i (1-\kappa_i^W)\right) \tag{A.31}$$

A.3.3 Augmented posterior distribution

Let $L_i^* = (r_i^{RD}, r_i^{LT}, \gamma_i, \kappa_i^{*,RD}, \kappa_i^{*,LT})$ be the latent, subject-specific variables in the model, and $Y^* = Y \cup L^* \cup \{\kappa_i^{RD}, \kappa_i^{LT}\}_{i=1}^N$. $(\kappa_i^{*,RD}, \kappa_i^{*,LT})$ are latent variables governing the process:

$$\kappa_i^W = \begin{cases} 1 & \text{if } \kappa_i^{*,W} \geq 0 \\ 0 & \text{otherwise} \end{cases}, \quad W \in \{RD, LT\} \tag{A.32}$$

As in (A.20) I use the following result to derive the joint posterior distribution of (θ, L^*), up to a normalizing constant:

$$\begin{aligned} p(\theta, L^* \mid Y) &\propto p(Y, L^* \mid \theta)p(\theta) = p(Y^* \mid \theta)p(\theta) && \text{(A.33)} \\ &= \prod_{\forall i} \left[p_i^{RD}(Y_i^* \mid \theta)p_i^{LT}(Y_i^* \mid \theta)\right] p(L_i^* \mid \theta) && \text{(A.34)} \\ &\quad \times p(\lambda_{RD})p(\lambda_{LT})p(\beta, \Sigma) \\ &\quad \times \prod_i \prod_{W \in \{RD, LT\}} I\left(\kappa_i^{*W}(2\kappa_i^W - 1) \geq 0\right) \end{aligned}$$

Comparing (A.20) and (A.34) reveals many similarities: firstly, the logit choice precision parameters $\lambda_R D$ and λ_{LT} appear in exactly the same way. Therefore the process for drawing these from their conditional posteriors is exactly the same; secondly, the latent variables Z^* appear enter into both equations in the same way, so drawing these will be identical, although one must now update the unconditional means, say μ_r, to account for the dependence between these parameter. One important difference is the final line of (A.34). This keeps track of the mapping between κ_i^* and κ_i.

The remainder of this section outlines the process for simulating the distribution in (A.34). This is achieved using a Gibbs sampler, a Markov chain Monte Carlo technique which requires that the conditional distributions of all variables can be simulated. I omit the derivation of the conditional posteriors when they are unchanged from the previous section.

Conditional distributions of κ_i^W and κ_i^{*W} These parameters are linked in that the sign of κ_i^{*W} fully determines κ_i^W. I draw from $p(\kappa_i^W, \kappa_i^{*W} \mid \theta, Y^*_{-(\kappa_i^W, \kappa_i^{*W})})$ by noting that:

$$p(\kappa_i^W, \kappa_i^{*W} \mid \theta, Y^*_{-(\kappa_i^W, \kappa_i^{*W})}) = p(\kappa_i^{*W} \mid \theta, Y^*_{-\kappa_i^{*W}}) p(\kappa_i^W \mid \theta, Y^*_{-(\kappa_i^W, \kappa_i^{*W})}) \tag{A.35}$$

I therefore draw κ_i^W first, then condition on this to draw κ_i^{*W}.

Drawing κ_i^W is identical to the process used in (A.23), however now ρ^W depends on subject characteristics, as well as the subject's other parameters. I now use the notation $\bar{\rho}_i^W$ to acknowledge that this parameter is subject-specific. Since L_i^* is multivariate normal, it follows that:

$$\bar{\rho}_i^W = p(\kappa_i^W = 1 \mid Y, X, \theta, L^*_{-\kappa_i^W, \kappa_i^{*W}}) \tag{A.36}$$

$$= p\left[\mathcal{N}\left(m_{\kappa^*W}, v_{\kappa^*W}\right) \geq 0\right] \tag{A.37}$$

$$= \Phi\left(\frac{m_{\kappa^*W}}{\sqrt{v_{\kappa^*W}}}\right) \tag{A.38}$$

$$\text{where: } m_{\kappa^*W} = X\beta_{\kappa^*W} + \Sigma_{\kappa^*W}\Sigma^{-1}_{-\kappa^*W}(L^*_{i-\kappa_i^{*W}} - \tilde{X}\tilde{\beta}_{-\kappa^*W}) \tag{A.39}$$

$$v_{\kappa^*W} = \sigma^2_{\kappa^*W} - \Sigma_{\kappa^*W}\Sigma^{-1}_{-\kappa^*W}\Sigma'_{\kappa^*W} \tag{A.40}$$

where $\sigma^2_{\kappa^*W}$ is the element of Σ corresponding to the variance of κ^{*W}, Σ_{κ^*W} is a $(1 \times (p-1))$ vector of covariances between κ_i^{*W} and other elements of L_i^*, and $\Sigma_{-\kappa^*W}$

is the $((p-1)\times(p-1))$ matrix left over after removing the row and column of Σ corresponding to κ^{*W}

The conditional posterior for κ_i^{*W} is proportional to:[6]

$$p(\kappa_i^{*W} \mid \theta, Y^*_{-\kappa_i^{*W}}) \propto p(L_i^* \mid \theta) I(\kappa_i^{*W}(2\kappa_i^W - 1) \geq 0) \tag{A.41}$$

Recognizing that the first term of (A.41) is the multivariate normal kernel, it follows that κ_i^{*W} conditional on θ and $Y^*_{-\kappa_i^{*W}}$ is distributed truncated normal:

$$\kappa_i^{*W} \mid \theta, Y^*_{-\kappa_i^{*W}} \sim \begin{cases} \mathcal{TN}_{(-\infty,0)}\left(m_{k^{*W}}, v_{k^{*W}}\right) & \text{if } \kappa_i^W = 0 \\ \mathcal{TN}_{[0,\infty)}\left(m_{k^{*W}}, v_{k^{*W}}\right) & \text{if } \kappa_i^W = 1 \end{cases} \tag{A.42}$$

Conditional distributions of other individual parameters r_i^{RD}, r_i^{LT}, and γ_i These parameters are drawn using the same Metropolis-Hastings method described in the previous section. As the parameters are correlated, the means and variances are adjusted to reflect this, as they are in (A.39) and (A.40).

Conditional distribution of model hyperparameters (β, Σ): For the most general model studied here, I assume that (transforms of) the individual parameters are draws from a joint normal distribution, that is:

$$L_i^* \mid \theta, X \equiv (r_i^{RD}, r_i^{LT}, \kappa_i^{*,RD}, \kappa_i^{*,LT}, \Phi^{-1}(\gamma_i)) \mid X \sim \mathcal{N}(X_i\beta, \Sigma) \tag{A.43}$$

where X is a $N \times k$ vector of subject characteristics (e.g. survey responses), β is $k \times p$, Σ is a $p \times p$ variance-covariance matrix, and $\Phi^{-1}(\cdot)$ is the inverse normal cdf.[7]

[6] Note that while $\kappa_i^{*W} \in Y_i^*$, κ_i^{*W} does not enter into $p_i^W(Y_i^* \mid \theta)$.

[7] Here, k is the number if explanatory variables used in X (including the constant term), and $p = |L_i^*|$ is the number of latent parameters used to model the data-generating process. $p = 5$ if parameters are allowed to vary between tasks, $p = 3$ if not.

Inspection of (A.20) yields that elements of β and Σ appear in the likelihood functions for both tasks, but only in the $p(Z_i \mid \theta)$ term. These parameters also appear in the prior term $p(\beta, \Sigma)$. Therefore, the conditional posterior for (β, Σ) reduces to:

$$
\begin{aligned}
& p(\beta, \Sigma \mid \lambda^{LT}, \lambda^{RD}, L^*) \\
& \propto p(\beta, \Sigma) \prod_{\forall i} \frac{|\Sigma|^{-\frac{1}{2}}}{(2\pi)^{\frac{k}{2}}} \exp\left[-\frac{1}{2}(L_i^* - X_i\beta)\Sigma^{-1}(L_i^* - X_i\beta)'\right] && \text{(A.44)} \\
& = p(\beta, \Sigma) \frac{|\Sigma|^{-\frac{N}{2}}}{(2\pi)^{\frac{Nk}{2}}} \exp\left[-\frac{1}{2}\sum_{\forall i}(L_i^* - X_i\beta)\Sigma^{-1}(L_i^* - X_i\beta)'\right] && \text{(A.45)}
\end{aligned}
$$

Now using the following priors:

$$
p(\beta_j) \sim \mathcal{N}(\underline{\beta}_j, \underline{V}_j) \quad j = 1, 2, \dots k \qquad \text{(A.46)}
$$

$$
H \equiv \Sigma^{-1} \sim \mathcal{W}(\underline{A}, \underline{v}) \qquad \text{(A.47)}
$$

(A.45) is proportional to:

$$
\begin{aligned}
& \prod_{j=1}^{|L_i^*|} \left[\frac{|\underline{V}_j|^{-\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}} \exp\left(-\frac{1}{2}(\beta_j - \underline{\beta}_j)' V_j^{-1} (\beta_j - \underline{\beta}_j)\right) \right] \\
& \times |H|^{\frac{\underline{v}-k-1}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}(\underline{A}^{-1}H)\right] && \text{(A.48)} \\
& \times |H|^{\frac{N}{2}} \exp\left[-\frac{1}{2}\sum_{\forall i}(L_i^* - X_i\beta)H(L_i^* - X_i\beta)'\right]
\end{aligned}
$$

where p is the number of explanatory variables contained in X. Inspection of (A.48) reveals that:

$$
\begin{aligned}
&p(H \mid \theta_{-H}, Y, L^*) \\
&\propto |H|^{\frac{\underline{v}-k-1}{2}} \exp\left[-\frac{1}{2}\text{tr}(\underline{A}^{-1}H)\right] |H|^{\frac{N}{2}} \exp\left[-\frac{1}{2}\sum_{\forall i}(L_i^* - X_i\beta)H(L_i^* - X_i\beta)'\right] && \text{(A.49)}\\
&= |H|^{\frac{N+\underline{v}-k-1}{2}} \exp\left[-\frac{1}{2}\text{tr}(\underline{A}^{-1}H) - \frac{1}{2}\sum_{\forall i}(L_i^* - X_i\beta)H(L_i^* - X_i\beta)'\right] && \text{(A.50)}\\
&= |H|^{\frac{N+\underline{v}-k-1}{2}} \exp\left[-\frac{1}{2}\text{tr}(\underline{A}^{-1}H) - \frac{1}{2}\sum_{\forall i}\text{tr}\left((L_i^* - X_i\beta)H(L_i^* - X_i\beta)'\right)\right] && \text{(A.51)}\\
&= |H|^{\frac{N+\underline{v}-k-1}{2}} \exp\left[-\frac{1}{2}\text{tr}(\underline{A}^{-1}H) - \frac{1}{2}\sum_{\forall i}\text{tr}\left((L_i^* - X_i\beta)'(L_i^* - X_i\beta)H\right)\right] && \text{(A.52)}\\
&= |H|^{\frac{N+\underline{v}-k-1}{2}} \exp\left\{-\frac{1}{2}\text{tr}\left[\left(\underline{A}^{-1} + \sum_{\forall i}(L_i^* - X_i\beta)'(L_i^* - X_i\beta)\right)H\right]\right\} && \text{(A.53)}
\end{aligned}
$$

noting that (A.53) is the Wishart kernel, it follows that the posterior conditional for H is:

$$
H \mid \theta_{-H}, Y, L^* \sim \mathcal{W}(\overline{A}, \overline{v}) \tag{A.54}
$$

$$
\text{where:} \quad \overline{A} = \left(\underline{A}^{-1} + \sum_{\forall i}(L_i^* - X_i\beta)'(L_i^* - X_i\beta)\right)^{-1}
$$

$$
\overline{v} = \underline{v} + N
$$

Going back to (A.48), the component of this proportional to β is:

$$
\exp\left[-\frac{1}{2}\sum_{\forall i}(L_i^* - X_i\beta)H(L_i^* - X_i\beta)'\right] \times \prod_{j=1}^{|L_i^*|} \exp\left(-\frac{1}{2}(\beta_j - \underline{\beta}_j)'V_j^{-1}(\beta_j - \underline{\beta}_j)\right) \tag{A.55}
$$

I make the following transformation of $\{L_i^*\}$ and relevant other variables from vector to matrix form:

$$L^* \equiv \begin{bmatrix} L_1'^* \\ L_2'^* \\ \vdots \\ L_N'^* \end{bmatrix} = \underbrace{\begin{bmatrix} I \otimes X_1 \\ I \otimes X_2 \\ \vdots \\ I \otimes X_N \end{bmatrix}}_{\tilde{X}} \underbrace{\text{vec}(\beta)}_{\tilde{\beta}} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix} \tag{A.56}$$

That is:

$$L^* \mid \theta, X \sim \mathcal{N}\left(\tilde{X}\tilde{\beta}, I \otimes \Sigma\right) \tag{A.57}$$

Therefore I can re-write (A.55) as:

$$p(\tilde{\beta} \mid \theta_{-\tilde{\beta}}, Y, L^*) \propto$$

$$\exp\left(-\frac{1}{2}(L^* - \tilde{X}\tilde{\beta})'(I \otimes \Sigma)^{-1}(L^* - \tilde{X}\tilde{\beta})\right) \exp\left(-\frac{1}{2}(\tilde{\beta} - \underline{\tilde{\beta}})'\underline{\tilde{V}}^{-1}(\tilde{\beta} - \underline{\tilde{\beta}})\right) \tag{A.58}$$

$$= \exp\left[-\frac{1}{2}\left((L^* - \tilde{X}\tilde{\beta})'(I \otimes \Sigma)^{-1}(L^* - \tilde{X}\tilde{\beta}) + (\tilde{\beta} - \underline{\tilde{\beta}})'\underline{\tilde{V}}^{-1}(\tilde{\beta} - \underline{\tilde{\beta}})\right)\right] \tag{A.59}$$

$$\propto \exp\left[-\frac{1}{2}\left(-2(\tilde{X}\tilde{\beta})'(I \otimes \Sigma)^{-1}L^* + (\tilde{X}\tilde{\beta})'(I \otimes \Sigma)^{-1}\tilde{X}\tilde{\beta} + \tilde{\beta}'\underline{\tilde{V}}^{-1}\tilde{\beta} - 2\tilde{\beta}'\underline{\tilde{V}}^{-1}\underline{\tilde{\beta}}\right)\right] \tag{A.60}$$

$$= \exp\left\{-\frac{1}{2}\left[\tilde{\beta}' \underbrace{\left(\tilde{X}'(I \otimes \Sigma)^{-1}\tilde{X} + \underline{\tilde{V}}^{-1}\right)}_{=A} \tilde{\beta} - 2\tilde{\beta}' \underbrace{\left(X'(I \otimes \Sigma)^{-1}L^* + \underline{\tilde{V}}^{-1}\underline{\tilde{\beta}}\right)}_{=B}\right]\right\} \tag{A.61}$$

$$= \exp\left\{-\frac{1}{2}\left[\tilde{\beta}'A\tilde{\beta} - 2\tilde{\beta}'AA^{-1}B\right]\right\} \tag{A.62}$$

$$\propto \exp\left\{-\frac{1}{2}\left[\tilde{\beta}'A\tilde{\beta} - \tilde{\beta}'AA^{-1}B - B'A'^{-1}A\tilde{\beta} + B'A'^{-1}AA^{-1}B\right]\right\} \tag{A.63}$$

$$= \exp\left\{-\frac{1}{2}\left[\left(\tilde{\beta} - A^{-1}B\right)' A \left(\tilde{\beta} - A^{-1}B\right)\right]\right\} \tag{A.64}$$

which is the multivariate normal kernel. Therefore:

$$\tilde{\beta} \mid \theta_{-\tilde{\beta}}, Y, L^* \sim \mathcal{N}\left(A^{-1}B, A^{-1}\right) \tag{A.65}$$

$$\text{where:} \quad A = \tilde{X}'(I \otimes \Sigma)^{-1}\tilde{X} + \underline{\tilde{V}}^{-1}$$

$$B = X'(I \otimes \Sigma)^{-1}L^* + \underline{\tilde{V}}^{-1}\underline{\tilde{\beta}}$$

A.4 Instructions

General Instructions

You are now taking part in an experiment. If you read the following instructions carefully, you can, depending on your and other participants' decisions, earn a considerable amount of money. It is therefore important that you take your time to understand the instructions. Please do not communicate with the other participants during the experiment. Should you have any questions, please raise your hand and an experimenter will come to you to answer your question in private.

The experiment consists of three different parts. In each part you will be asked to make one or more decisions. You will have to make your decisions without knowing other participants' decisions in the previous parts. Note further that the other participants will also not know your decisions.

You will receive specific instructions for each part, only once the previous part has been completed.

Once the experiment has finished, your earnings will be paid to you in cash. You will then be allowed to leave the lab and no one except you will know either your earnings or your decisions.

Note that, at the end of the experiment, the computer will randomly determine which of your decisions will determine your earnings. You will receive more information about this in the instructions specific to that part.

For the rest of these instructions, and for all tasks in the experiment, payments will be expressed in experimental dollars, which will be converted to US dollars at the end of the experiment at an exchange rate of:

1 Experimental Dollar = US$0.02

Before beginning each part of the experiment, you will be given some instructions. After reading the instructions, you will be given some questions to test your understanding of these instructions. For every correct answer, you will have 20 experimental dollars added to your earnings.

Instructions Part 1

In Part 1, you will be placed in a group of four, including yourself and three other individuals. This group will remain the same for all of Part 1. Part 1 consists of 20 periods. The computer will randomly select five of these periods to be added to your earnings. Since you will not know which of the periods will be added to your earnings until the end of the experiment, you should make your decisions in each period as if they count for payment.

In each period, you will be randomly paired with another individual in your group. From now on we will refer to this individual as "the other individual". As the random parings occur each period, you will not know whether or not you are performing this task paired with the same person or another person from one period to the next.

Your task in each period is to choose between Action A and Action B, and also to choose between Action C and Action D. At the same time, the other individual will also be choosing between these same two actions. This task is shown below:

		The other's action			
		A		B	
Your action	○ A	You earn: a	Other earns: a	You earn: a	Other earns: x
	○ B	You earn: x	Other earns: a	You earn: w	Other earns: w

		The other's action			
		C		D	
Your action	○ C	You earn: c	Other earns: c	You earn: c	Other earns: Y
	○ D	You earn: y	Other earns: c	You earn: z	Other earns: z

During the experiment, the letters denoting earnings will be replaced with numbers, so you will know how much you could earn from taking each action

The entries in each of these tables show the earnings that result from the combination of your actions and the other individual's actions. The red entry on the left in each cell shows your earnings, and the blue entry on the right shows the other individual's earnings. For example, for the top table, if you choose Action B and the other individual chooses Action A, then you earn *x* experimental dollars, and the other individual earns *a* experimental dollars. Likewise, for the bottom table, if you choose Action C and the other individual chooses Action C, then you earn *c* experimental dollars, and the other individual earns *c* experimental dollars. If this period is chosen for payment, then in total you would earn *x+c* experimental dollars, and the other individual would earn *a+c* experimental dollars.

In all periods except the first, you will be shown a table summarizing the actions taken by all of the individuals in your group in all previous periods. An example of this table (with the entries blanked out), is shown below:

Period	A & C	A & D	B & C	B & D	My choices	Other's choices
1						
2						
3						
4						
5						

The leftmost column shows the period number. The next four columns show the number of individuals in your group, including yourself, that chose the four possible combinations of actions in that period (if the number is zero, it will be blanked out to make reading the rest of the table easier). The second column from the right lists the choices you made in previous periods, and the rightmost column shows the actions that the other individual chose. Remember that this other individual may change between periods because the pairings change randomly in every period.

At the end of the experiment, you will be shown the results of the periods chosen for payment, and the amounts that will be added to your earnings.

What information do the three other participants in my group have?

All participants have received the same instructions about the task in Part 1. During Part 1 they will be shown the same table describing the task. They will see the same table summarizing actions taken by the group in the past, except that the two rightmost columns of this table will show the action that *they* took, and the action that *their* "other individual" took for that period.

By the time you start Part 1, all participants will have answered the questionnaire, which will appear on your screen.

Do you have any questions? If so, please raise your hand and an experimenter will come to you and answer your questions in private. Otherwise, please complete the questionnaire on your computer screen. Remember, 20 experimental dollars will be added to your earnings for every correct answer in this questionnaire.

Instructions for Part 2

Part 2 of the experiment consists of 4 periods. In each period you will start with 100 tokens, each token is worth one experimental dollar. Your task is to decide the number of tokens to invest in a lottery described below.

An example of your task in a period (which shall not be used in the experiment) is presented below:

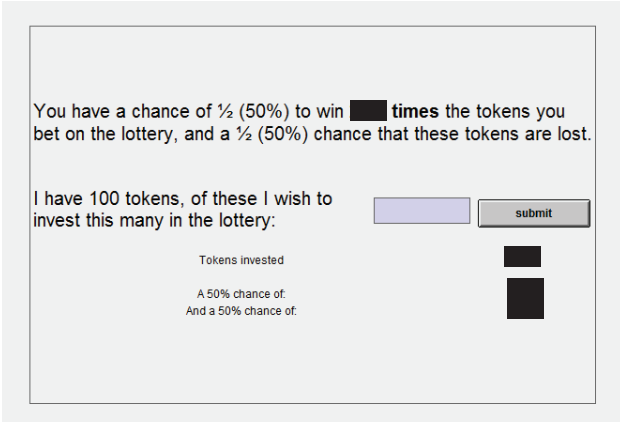
> *You have a chance of ½ (50%) to win* ***3.00 times*** *the tokens you bet on the lottery, and a ½ (50%) chance that these tokens are lost.*

Hence, your tokens at the end of this period are determined as follows. If you decided to put amount *Y* tokens (and keep *100 – Y* tokens out of the lottery), then you will have at least *100 – Y* tokens at the end of this period. In addition to this, if the lottery is successful you will have *3.00 Y* additional tokens at the end of the period. In other words, if the lottery is unsuccessful you will end the period with *100-Y* tokens, and if the lottery is successful, you will end the period with *100-Y +3.00Y* tokens

You can think of the outcome of this lottery as the result of tossing a fair coin. You receive the additional *3.00Y* tokens whenever the coin comes up heads.

Note that the prize of the lottery will change between periods, so please read each lottery description carefully before making your decisions.

When you enter the number of tokens you would like to invest, we first require that you "submit" this number to the computer. Submitting this number will calculate the tokens that you will have at the end of the period in the case that the lottery is not successful, and the case that it is successful. An example of the results of this calculation is presented in the figure below (the numbers are blanked out):



Clicking “submit” will not lock in your answer, so feel free to experiment with different numbers until you find one that you like. When you wish to lock in your answer, click “OK”.

Will other participants learn about your decisions in this Part?

No, other participants will not be informed of your decisions in this part.

How will your earnings be determined for this Part?

One decision in either Part 2 or Part 3 will be selected for payment. At the end of the experiment, the computer will randomly select either Part 2 or Part 3 for payment. Each part is equally likely to be selected. If Part 2 is selected, then the computer will randomly select one of the four periods in this Part. Each period is equally likely to be selected. For the selected period, one experimental dollar will be added to your final earnings for every token that you have at the end of this period. Since you will not know which of the period will be added to your earnings until the end of the experiment, you should make your decisions in each period as if they count for payment.

When will you find out about the outcomes of your decisions in this Part?

If this part is selected for payment, then you will be informed of the outcome of the lottery that was selected for payment at the end of the experiment, after all decisions have been made.

Do you have any questions? If so, please raise your hand and an experimenter will come to you and answer your questions in private. Otherwise, please complete the questionnaire on your computer screen. Remember, 20 experimental dollars will be added to your earnings for every correct answer in this questionnaire. Payment for correctness will occur irrespective of whether Part 2 or 3 is selected for payment.

Instructions for Part 3

Part 3 of the experiment consists of 4 periods. In each period you will start with 100 red tokens, and 100 blue tokens, each token is worth one experimental dollar. Your task is to decide the number of red and blue tokens to invest in the "red lottery" and the "blue lottery", respectively, described below.

An example of your task in a period (which shall not be used in the experiment) is presented below:

RED LOTTERY	BLUE LOTTERY
You have a chance of ½ (50%) to win ***2.25 times*** *the tokens you bet on the lottery, and a ½ (50%) chance that these tokens are lost.*	*You have a chance of ½ (50%) to win* ***2.25 times*** *the tokens you bet on the lottery, and a ½ (50%) chance that these tokens are lost.*

Hence, your tokens at the end of this period are determined as follows. If you decided to put *R* tokens in the red lottery (and keep *100 – R* tokens out of the lottery), and put *B* tokens in the blue lottery (and keep *100 – B* tokens out of the lottery), then you will have at least *100 – R* red tokens and *100 – B* blue tokens at the end of this period. In addition to this, if the red lottery is successful you will have *2.25 R* additional red tokens at the end of the period, and if the blue lottery is successful, you will have *2.25 B* additional blue tokens at the end of the period.

You can think of the outcomes of these lotteries as the result of tossing two fair coins, a "red coin" and a "blue coin". You receive the additional *2.25 R* red tokens whenever the red coin comes up heads, and you receive the additional *2.25 B* blue tokens whenever the blue coin comes up heads.

Note that the prizes of the red and blue lotteries will change between periods, so please read each lottery description carefully before making your decisions.

As in Part 2, we require that you "submit" your answers for both lotteries before clicking "OK". Again, submitting your answer will calculate the tokens that you will have at the end of the period in the case that the lottery is not successful, and the case that it is successful. Clicking "submit" will not lock in your answer, so feel free to experiment with different numbers until you find one that you like. When you wish to lock in your answer, click "OK".

Will other participants learn about your decisions in this Part?

No, other participants will not be informed of your decisions in this part.

How will your earnings be determined for this Part?

One decision in either Part 2 or Part 3 will be selected for payment. At the end of the experiment, the computer will randomly select either Part 2 or Part 3 for payment (each part is equally likely to be selected). If Part 3 is selected, then the computer will randomly select one of the four periods in this Part (each period is equally likely to be selected), and one experimental dollar will be added to your final earnings for every token (both red and blue) that you have at the end of this period. Since you will not know which of the periods will be added to your earnings until the end of the experiment, you should make your decisions in each period as if they count for payment.

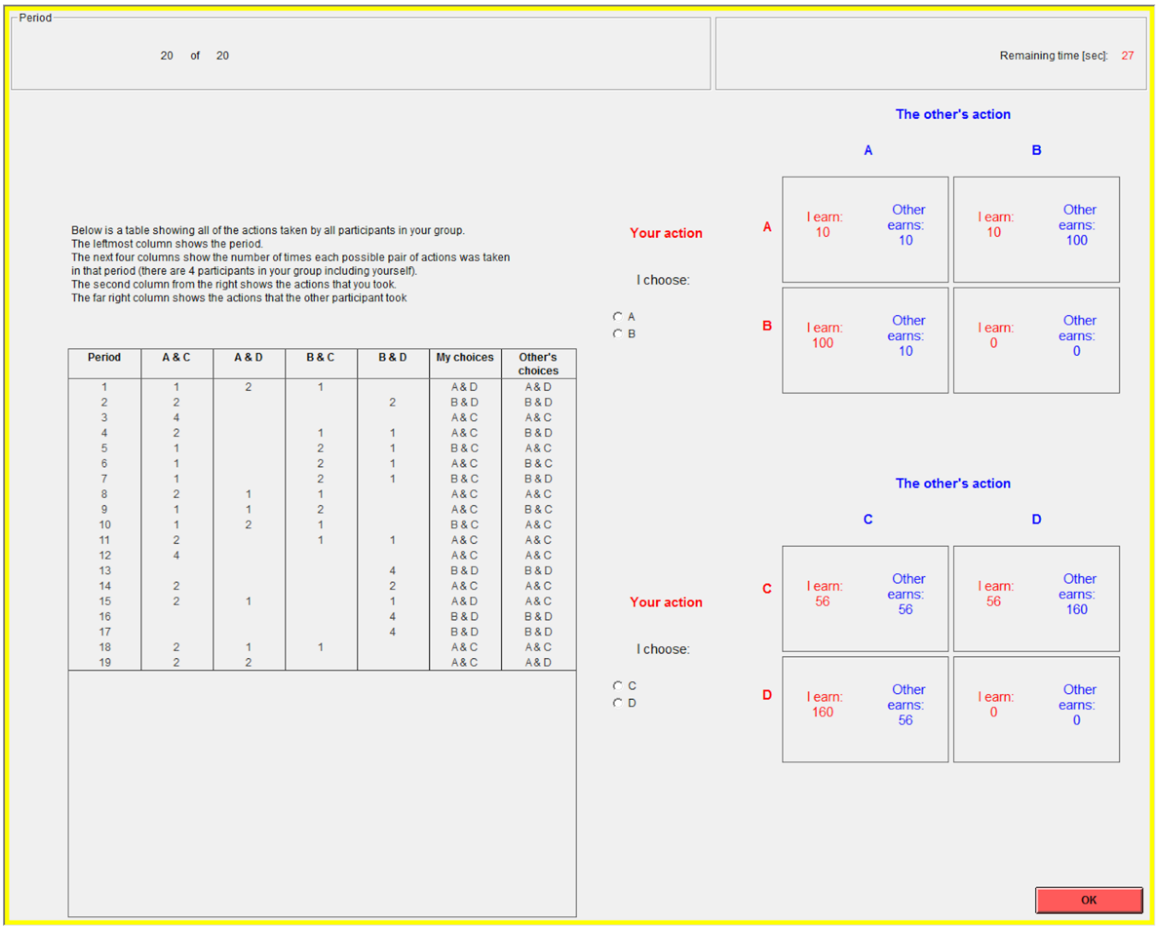
When will you find out about the outcomes of your decisions in this Part?

If this part is selected for payment, then you will be informed of the outcome of the lotteries that were selected for payment at the end of the experiment, after all decisions have been made.

Do you have any questions? If so, please raise your hand and an experimenter will come to you and answer your questions in private. Otherwise, please complete the questionnaire on your computer screen. Remember, 20 experimental dollars will be added to your earnings for every correct answer in this questionnaire. Payment for correctness will occur irrespective of whether Part 2 or 3 is selected for payment.

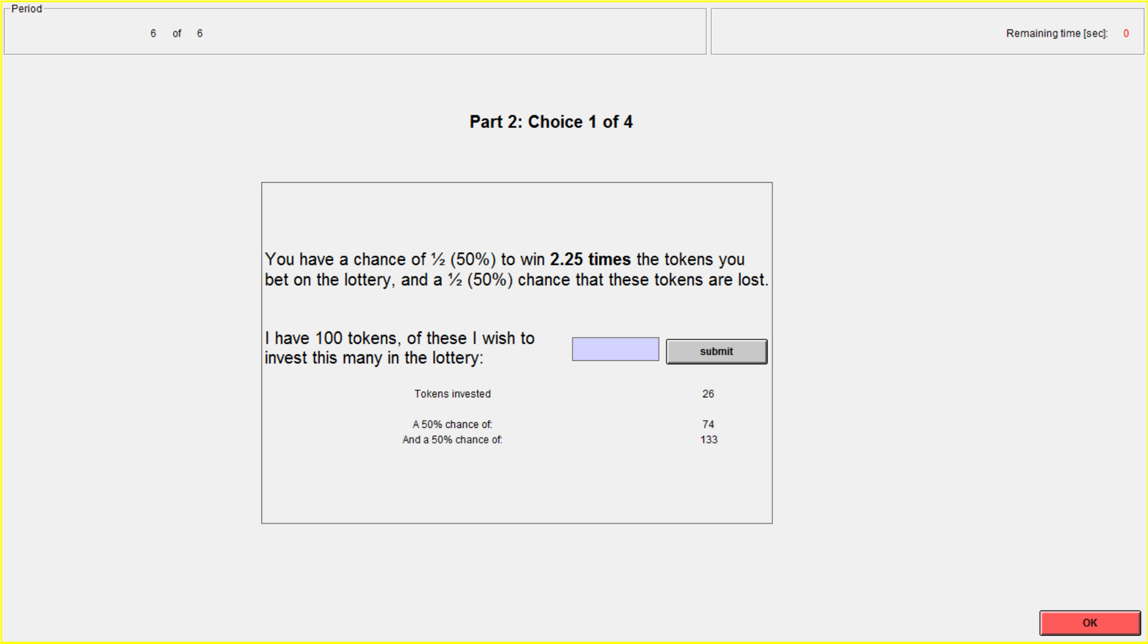
A.5 Screenshots

Roommates' Dilemma (Treatment 1)

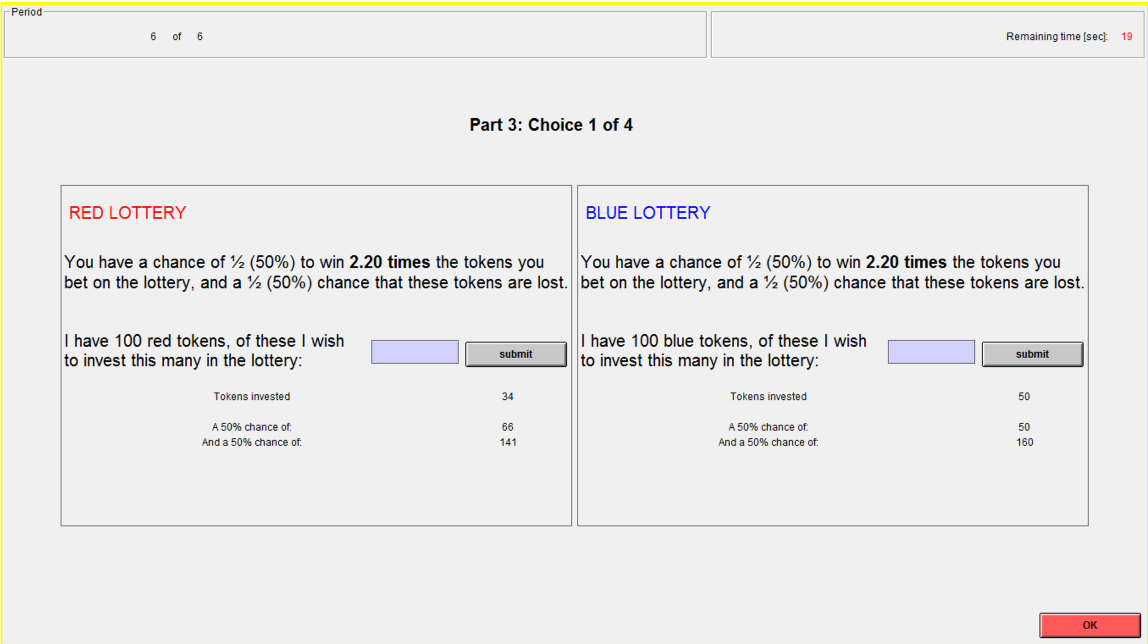


Period	A & C	A & D	B & C	B & D	My choices	Other's choices
1	1	2	1		A & D	A & D
2	2			2	B & D	B & D
3	4				A & C	A & C
4	2		1	1	A & C	B & D
5	1		2	1	B & C	A & C
6	1		2	1	A & C	B & C
7	1		2	1	B & C	B & D
8	2	1	1		A & C	A & C
9	1	1	2		A & C	B & C
10	1	2	1		B & C	A & C
11	2		1	1	A & C	A & C
12	4				A & C	A & C
13				4	B & D	B & D
14	2			2	A & C	A & C
15	2	1		1	A & D	A & C
16				4	B & D	B & D
17				4	B & D	B & D
18	2	1	1		A & C	A & C
19	2	2			A & C	A & D

Lottery task part 1



Lottery task part 2



B. APPENDIX FOR: MIXTURE MODELS OF BEHAVIOR AND NUISANCE PARAMETERS: A SEMI-PARAMETRIC BAYESIAN APPROACH

B.1 Additional notes on Andreoni and Vesterlund (2001)

B.1.1 Derivatoin of optimal choice for each type

For the PC type:

$$t^*_{i,k} = \arg\max_{t\in[0,1]} \min\{\theta_i^{PC} t, (1-t)/p_k\} \tag{B.1}$$

$$\theta_i^{PC} t^*_{i,k} p_k = (1 - t^*_{i,k}) \tag{B.2}$$

$$t^*_{i,k} = \frac{1}{p_k \theta_i^{PC} + 1} \tag{B.3}$$

For the PS type:

$$t^*_{i,k} = \arg\max_{t\in[0,1]} t + \theta_i^{PS}(1-t)/p_k \tag{B.4}$$

$$t^*_{i,k} = \begin{cases} 1 & \text{if } p_k \geq \theta_i^{PC} \\ 0 & \text{if } p_k < \theta_i^{PC} \end{cases} \tag{B.5}$$

here we assume that when indifferent, the PS type keeps the endowment.

B.1.2 Calculation of lower bound on nuisance parameters

For the PC type, we require that when $p_k = 1$ the lower bound on θ^{PC} havs a prediction that differs by at least ϵ from the selfish type:

$$\inf_{\theta^{PC}} \|1 - t^*_{i,p=1}\| \geq \epsilon \tag{B.6}$$

$$1 - \frac{1}{\theta^{PC} + 1} \geq \epsilon \tag{B.7}$$

$$\theta^{PC} \geq \epsilon\theta^{PC} + \epsilon \tag{B.8}$$

$$\theta^{PC} \geq \frac{\epsilon}{1 - \epsilon} \tag{B.9}$$

For the PS type, we require that for at least one treatment the lower bound type passes all of the endowment, which according to (B.5) means that:

$$\theta^{PS} > \min_k (p_k) = \frac{1}{3} \tag{B.10}$$

C. APPENDIX FOR: HOSPITAL-INSURER BARGAINING POWER AND NEGOTIATED RATES

C.1 Understanding Variation in Fee For Service Payments

In this section, we explore the variation in FFS payments, and some of the issues with using charge master prices as a measure of services. We assume that FFS payments are negotiated as a fraction of charges. If all variation in FFS payments were explained by this negotiated fraction, then a regression of log(payments) against log(charges) should result in a coefficient of 1 on log(charges), and and R^2 of 1. Specifically, estimating the model:

$$\log(\text{paid}_j) = \beta_0 + \beta_1 \log(\text{charged}_j) + \epsilon_j \tag{C.1}$$

for all patients j who were admitted at hospital h with insurance i should result in estimates of $\hat{\beta}_1 = 1$ and $R^2 = 1$. Setting $\epsilon_j = 0 \ \forall j$ yields:

$$\log\left(\frac{\text{paid}_j}{\text{charged}_j}\right) = \beta_0 \tag{C.2}$$

$$\rho_{hi} = \exp(\beta_0) \tag{C.3}$$

These ρ_{hi}s tell us the fraction of charges that a homogenized patient would pay for a particular hospital-insurer combination. As such they allow us to compare relative prices between hospital-insurer pairs.

We begin with Table C.1 by exploring the variation in FFS payments and charges due to characteristics such as hospital type, location, and insurance and hospital market share. These regressions are used to motivate our analysis based on a normalization of charges, rather than charges themselves.

Before interpreting the coefficients on these regressions, it is worthwhile noting that, given a long enough time horizon, all of these market characteristics are endogenous. While we expect hospital location, teaching status, and the presence of an emergency room to be choice variables for the hospital only in the *very* long run, and hence fixed for the purposes of a particular negotiation, market share clearly varies directly with the negotiation process. We therefore invite the reader to be cautious in interpreting the coefficients on market shares as causal.[1]

We have four specifications. The first specification examines payments and charges on a per-patient basis. The second specification explains payments considering illness heterogeneity by interacting illness weight and hospital type. The third specification explains charges again by considering illness heterogeneity by interacting illness weight and hospital type. The final specification uses normalized charges to explain payments.

In Table C.1 column (1), we regress log(payments) against log(charges), hospital type indicators, region indicators, and measures of market share of hospitals and insurers. For FFS payments, we expect the coefficient on log(charges) to be equal to one. While we reject this hypothesis at the 5% level, the coefficient of 0.940 indicates a close to linear relationship between payments and charges.[2]

We next turn our attention to hospital type. Of some concern in column (1) is the negative and significant coefficient on the academic medical center (AMC) indicator. The comparison group of hospital type is community hospitals, which we think of as being highly substitutable. We predicted community hospitals would have limited bargaining power. AMCs are high-quality specialized hospitals, so we expected this

[1]Specifically, we expect hospital (insurer) market share to be correlated with unobservable hospital quality (insurance policy) characteristics that are important in the negotiating process. For example, we expect higher quality hospitals to command both (i) greater market shares and (ii) greater negotiated rates. As quality is unobserved and therefore an omitted variable, this would mean that the estimated coefficient on market share would over-state the effect of having a greater market share. Hence the estimates are an upper bound of the causal affect of a hospital or insurer commanding a greater market share in the negotiation.

[2]One thing to observe is that this suggests that patients with higher charges may pay a slightly smaller fraction of the bill than patients with lower charges. This may be attributable to budget constraints. If patients pay 10% of their bills through coinsurance, patients with bills in the upper tail of the distribution may struggle to pay even the coinsurance fraction of their bill.

Table C.1.: Relationship between payments and charges for Fee For Service payments. All regressions control for sex, race, ethnicity, and age (polynomial). Hospitals classified as "other specialty" have been omitted from these regressions.

	(1) log(paid)	(2) log(paid)	(3) log(charged)	(4) log(paid)
log(charged)	0.937***			
	(0.0109)			
Academic MC	-0.481***	0.460***	0.696***	0.148***
	(0.0342)	(0.0365)	(0.0277)	(0.0332)
Community DSH	0.0793*	0.122**	0.0361	0.160***
	(0.0352)	(0.0391)	(0.0297)	(0.0357)
Teaching	-0.0182	0.183***	0.0323	0.0719*
	(0.0293)	(0.0331)	(0.0251)	(0.0292)
TraumaCenters	0.161***	0.135***	0.0400	0.0702*
	(0.0316)	(0.0298)	(0.0226)	(0.0316)
Central Mass	-0.0568	-0.237***	-0.145***	-0.250***
	(0.0423)	(0.0473)	(0.0359)	(0.0425)
North East Mass	-0.211***	-0.104*	-0.00933	-0.145***
	(0.0397)	(0.0426)	(0.0323)	(0.0399)
Metro Boston	0.497***	0.0328	-0.0483	0.528***
	(0.0382)	(0.0381)	(0.0289)	(0.0388)
South Eastern Mass	0.163***	0.00504	-0.220***	-0.126**
	(0.0424)	(0.0460)	(0.0349)	(0.0428)
Log hospital market share (DRG weight)	0.143***			0.256***
	(0.0116)			(0.0119)
Log insurer market share (DRG weight)	-0.131***			-0.124***
	(0.00451)			(0.00451)
Academic MC × log(DRGweight)		0.266***	0.373***	
		(0.0251)	(0.0190)	
Community × log(DRGweight)		0.294***	0.382***	
		(0.0296)	(0.0225)	
Community DSH × log(DRGweight)		0.269***	0.401***	
		(0.0301)	(0.0228)	
Teaching × log(DRGweight)		0.358***	0.414***	
		(0.0335)	(0.0254)	
log(Imputed weight)				1.044***
				(0.0118)
Constant	-0.705	7.557***	8.116***	8.514***
	(0.520)	(0.465)	(0.352)	(0.526)
Observations	43889	9797	9797	43462
R^2	0.220	0.167	0.403	0.228

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

coefficient to be positive. The reason for this unexpected sign is investigated in columns (2) and (3).

Column (2) in Table C.1 regresses log payments against various controls for market characteristics, and investigates the relationship between DRG weights (which capture illness heterogeneity) and payments.[3] One can interpret the coefficients on

[3] Here we include $\log(\text{weight}_j)$ interacted with all hospital type indicators to allow for more flexibility in this relationship. Typically in specifications with interactions, one includes a 'linear' term for the interacted variable alone. In this specification, however, since we interact weight with every hospital type, to include weight (uninteracted), we would have to omit the effect of an increase in weight on a particular hospital type.

the hospital type indicators as (approximately) the fraction markup over a community hospital for a DRG weight of one. Note here that academic medical centers are estimated to receive $1.57 = \exp(0.454)$ times the payment compared to a community hospital for a patient with a DRG weight of 1. This coefficient, unlike the corresponding coefficient in column (1), *is* in line with our expectations that AMCs are the high-quality hospitals, and hence command higher prices.

The coefficient `AMC` $\times$ `Log weight` of .267 should be interpreted as follows. For an increase in DRG weight of 1, payments would increase by a factor of $\exp(0.267) = 1.306$, or 30.6%.

Column (3) changes the dependent variable in the regression, and regresses log(charges) against the same regressors as column (2). This estimation highlights the discrepancy noted in column (1) where the AMC indicator's coefficient was expected to have a positive sign. The coefficient on AMC in column (3) is significantly positive, indicating that AMC's *charges* are uniformly higher. Therefore, an insurer with the same negotiated rate (as a fraction of charges) at a community hospital and an AMC would be charged more at the AMC.

Notice, though, when we compare columns (2) and (3), that `AMC` have much higher charges than the community hospitals. AMCs have much higher prices than their community counterparts, but they only get payments that are a fraction higher.

This observation motivates the normalization of charges. We normalize hospital charges by creating a hospital charge index which compares each patient's charges to estimated charges at that hospital for a DRG of 1. This procedure is described in detail in Section C.2.2.

We call this variable `log(Imputed Weight)` and use it instead of raw charges in the same specification as column (1). The results of this are reported in column (4) of Table C.1. Here we see that AMCs command payments of about $1.155 = \exp(0.143)$ times higher than community hospitals, and an additional 7% if they have a trauma center. When we compare our results to Ho (2008), who estimates 'star' hospitals to have markups of 25%, we find similar results (about 23%). We also include log market share by DRG weight, and estimate that a 1% increase in market share of a hospital is associated with a 0.257% increase in negotiated payments, and a 1% increase in market share of an insurer is associated with a 0.127% decline in negotiated payments. Again, due to suspected correlation with unobservable quality attributes, we consider these value to be upper bounds on the causal effects of these variables in the negotiation process.

C.2 Procedure for calculating negotiated rates

Our goal is to estimate the negotiated rates between hospitals and insurers in the market for health care. To this end, we must estimate:

1. An index of relative hospital charges. Specifically, for each hospital we estimate the amount that a patient with a DRG weight of 1 would be charged.
2. Negotiated rates for each hospital-insurer pair (i.e. how much is paid for services for a normalized patient[4], and how this varies for each hospital-insurer pair?)

The remainder of this section outlines the procedure.

C.2.1 Index of relative hospital charges

In Section C.1,we find a patient's charges varies greatly by hospital, even after normalizing differences in DRG weight.[5] As we aim to estimate negotiated rates as a fraction of charges, we wish to control for this variation so that a normalized "charge" buys the same quantity of services at each hospital.[6] As a motivation for this normalization, consider two hospitals, one with low costs and charges and one

[4]Normalized in this context means for a patient with a DRG weight of 1

[5]For example, we would imagine hospitals at the top of the referral network to treat sicker patients. When we say we are controlling for DRG weights, what we mean is that we are comparing charges per DRG weight. Thus we can compare charges as if patients were homogeneous and each had an average illness.

[6]One can roughly think of this as the ratio of actual charges to the charges a patient would receive if they had a weight of 1 (i.e., charges for an average patient) at that hospital. Precise details on our methodology for constructing these normalized charges are below

with high costs and high charges. $5,000 of charges at Hospital L(ow) would buy more services than $5,000 at Hospital H(igh). Our normalization allows us to look at percentage changes in service provision relative to a standard unit of service.[7]

We first calculate an index of hospital charges that is equal to the estimated (log) charges that a patient with a DRG weight of 1 would receive. To achieve this, we estimate the following equation for every hospital twice, once for FFS and once for DRG payments:

$$\log(\text{charged}_j) = \beta_{0,h} + \beta_{1,h} \log(\text{weight}_j) + \epsilon_j \tag{C.4}$$

where weight_j is patient j's DRG weight. Taking the conditional expectation of (C.4) for a weight of 1 yields:

$$E\left[\log(\text{charged}_j) \mid \text{weight}_j = 1\right] = \beta_{0,h} \tag{C.5}$$

Using the estimated intercept from (C.4), we then construct:

$$\log(\text{normalized\_charges}_j) = \log(\text{charged}_j) - \hat{\beta}_{0,h} \tag{C.6}$$

That is, if two patients at different hospitals have the same normalized charges, (say 1.5) we expect that they each receive 1.5 times as many services as a normalized patient at their hospital. the same quantity of services, but their (un-normalized) charges could be different.[8] Furthermore in expectation, $\log(\text{normalized\_charges}_j) = 0$

[7]A standard unit of service is one with a DRG weight of 1, and is not dependent on patient case mix.

[8]Because this specification is in logs, this is more clear if you consider 2 patients both of who receive 2 times the average service provision at two different hospitals. Both would have a normalized charge of 2. At hospital A, average charges may be $8,000. At hospital B, average charges may be $10,000. The patient at Hospital A would receive $16,000 of services and at Hospital b would receive $20,000, but each receives twice the average, so they each have a normalized charge of 2. In this way, normalized charges tells us the relative amount of service consumption at each hospital without explicitly comparing dollar values.

corresponds to the quantity of services that would be provided for a DRG weight of 1.[9]

C.2.2 Procedure for estimating negotiated rates for FFS payments

Fee-for-service (FFS) payment arrangements are negotiated payments between a hospital and an insurer specifying either amounts to be paid for each service provided by the hospital, or the fraction of the hospital's charges that will be paid. Due to the availability of data, we make the simplifying assumption that all FFS contracts are a negotiated fraction of charges, noting that the alternative would be a negotiated price vector for services, say $\mathbf{p}_{hi}$. If $\mathbf{p}_{hi} = \rho_{hi}\mathbf{c}_h$ for some $\rho_{hi} \in (0, 1)$, where $\mathbf{c}_h$ is hospital h's charge master prices, then these types of arrangements are identical.

We seek to estimate the fraction of charges that each insurer pays to each hospital. Additionally, noting that charges for the same services vary greatly between hospitals, we also estimate the amount paid by an insurer for one unit of normalized payments.

For a FFS arrangement, the relationship between charges and payments should be:

$$\text{paid}_j = \rho_{hi}\text{charged}_j, \quad \iff \quad \log(\text{paid}_j) = \log(\rho_{hi}) + \log(\text{charged}_j) \tag{C.7}$$

where ρ_{hi} is the fraction of charges paid by insurer i at hospital h. We estimate (C.7) by regressing log(payments) against log(charges) for all patients in a hospital-insurer pair:

$$\log(\text{paid}_j) = \beta_{0,hi} + \beta_{1,hi}\log(\text{charged}_j) + \epsilon_j \tag{C.8}$$

Hence, our estimated for negotiated rates as a fraction of services are:

$$\hat{\rho}_{hi} = \exp(\hat{\beta}_{0,hi}) \tag{C.9}$$

[9]and for small x, $\log(\text{normalized\_charges}_j) = x$ corresponds to a DRG weight of approximately $1+x$.

Note that between (C.7) and (C.8) we allow for $\beta_{1,hi} \neq 1$, that payments are not directly proportional to charges.

We also estimate the amount paid for normalized charges using the regression equation:

$$\log(\text{paid}_j) = \beta_{0,hi} + \beta_{1,hi} \log(\text{normalized\_charges}_j) + \epsilon_j \tag{C.10}$$

again, allowing for a nonlinear relationship between payments and normalized charges through estimating (rather than imposing a value on) $\beta_{1,hi}$. Our estimate of the amount paid for a unit of normalized charges is therefore:

$$\hat{\gamma}_{hi} = \exp(\beta_{0,hi}) \tag{C.11}$$

Note that the units of $\hat{\gamma}_{hi}$ are in US dollars, while $\hat{\rho}_{hi}$ is unit-free.

C.3 Estimates of negotiated rates

	(1) Aetna	(2) Blue Cross	(3) Cigna	(4) Fallon	(5) United	(6) Tufts	(7) ALL
Anna Jaques	0	1539.2	0	0	4072.7	0	2158.2
	(.)	(312.2)	(.)	(.)	(1424.0)	(.)	(400.4)
Baystate MC	0	3320.1	7786.1	4853.6	7792.0	0	4794.6
	(.)	(210.6)	(906.8)	(617.0)	(1052.4)	(.)	(245.9)
Baystate Franklin MC	0	2264.1	0	0	0	0	4061.5
	(.)	(486.9)	(.)	(.)	(.)	(.)	(574.0)
Baystate Mary Lane	0	2639.7	0	0	0	0	4301.0
	(.)	(883.1)	(.)	(.)	(.)	(.)	(979.3)
Brigham and Women's	22001.1	7559.8	19721.4	10682.3	11954.0	13640.7	13279.2
	(296.5)	(193.2)	(813.6)	(958.8)	(1003.1)	(245.6)	(163.9)
Signature Healthcare Brockton	0	2186.7	0	0	5639.6	7102.1	3248.0
	(.)	(215.4)	(.)	(.)	(1446.4)	(652.9)	(274.4)
Cape Cod	10620.6	3770.7	0	0	9507.6	9897.4	6300.0
	(887.0)	(298.7)	(.)	(.)	(1657.1)	(666.1)	(337.7)
Falmouth	0	4130.9	0	0	0	12880.7	7385.6
	(.)	(499.2)	(.)	(.)	(.)	(1029.8)	(540.3)
Steward Norwood	0	0	0	0	5107.2	7106.3	5924.0
	(.)	(.)	(.)	(.)	(1362.5)	(455.0)	(392.3)
Steward Carney	0	1247.1	0	0	5972.2	5910.0	3529.9
	(.)	(452.6)	(.)	(.)	(1413.2)	(942.0)	(469.2)
Boston Children's	14984.2	4115.8	13830.8	0	0	16476.6	7753.3
	(648.9)	(234.4)	(1191.8)	(.)	(.)	(746.6)	(276.6)
Cooley Dickinson	0	0	0	2672.1	0	6189.5	4331.9
	(.)	(.)	(.)	(647.1)	(.)	(737.2)	(557.2)
Beth Israel -Needham	0	1332.6	0	0	0	0	3510.8
	(.)	(709.7)	(.)	(.)	(.)	(.)	(819.3)
Emerson	0	3645.1	0	0	5333.5	8707.9	5640.4
	(.)	(318.1)	(.)	(.)	(1548.1)	(504.2)	(328.7)
Brigham and Women's-Faulkner	0	3258.9	0	0	8358.4	8408.0	5785.9
	(.)	(360.5)	(.)	(.)	(1838.4)	(680.1)	(392.8)
Harrington Memorial	0	2071.9	0	3743.8	0	0	3609.1
	(.)	(448.9)	(.)	(575.0)	(.)	(.)	(438.0)
Health Alliance	0	0	0	3560.8	7307.0	0	4487.0
	(.)	(.)	(.)	(520.7)	(1711.4)	(.)	(495.2)
Heywood	0	1933.4	0	2725.0	0	6122.7	3018.7
	(.)	(333.0)	(.)	(633.2)	(.)	(793.8)	(364.6)
Steward Holy Family	0	2380.4	0	0	7229.3	7645.2	5250.3
	(.)	(352.1)	(.)	(.)	(960.1)	(557.3)	(324.5)
Holyoke MC	0	2003.5	0	0	6521.8	0	3415.8
	(.)	(421.6)	(.)	(.)	(1815.3)	(.)	(489.6)

Beth Israel -Plymouth	0	2906.4	0	0	5173.5	7347.6	4219.3
	(.)	(291.3)	(.)	(.)	(1862.4)	(620.3)	(344.7)
Lawrence General	0	0	0	1702.6	0	7638.9	5907.0
	(.)	(.)	(.)	(886.1)	(.)	(580.6)	(523.5)
Lowell General	0	2908.2	0	0	5985.2	7898.8	4998.2
	(.)	(215.4)	(.)	(.)	(963.4)	(376.3)	(230.6)
Massachusetts Eye and Ear	0	0	0	0	0	5013.7	4743.4
	(.)	(.)	(.)	(.)	(.)	(1165.6)	(857.7)
Massachusetts General	17153.6	2823.5	14790.3	0	9857.6	12197.0	8859.3
	(323.9)	(157.6)	(892.8)	(.)	(692.3)	(233.0)	(148.4)
Milford Regional MC	0	2541.9	0	3900.3	5065.3	8892.4	4184.2
	(.)	(366.5)	(.)	(690.7)	(1771.5)	(824.2)	(372.0)
Beth Israel - Milton	0	1964.2	0	0	4480.5	5127.1	3493.5
	(.)	(445.2)	(.)	(.)	(1862.4)	(695.1)	(458.9)
Morton	0	1839.7	0	0	5710.9	5596.8	3059.0
	(.)	(316.7)	(.)	(.)	(1674.6)	(873.4)	(375.0)
Mount Auburn Hospital	0	4995.8	0	0	7668.5	0	5558.3
	(.)	(360.5)	(.)	(.)	(1771.5)	(.)	(456.2)
New England Baptist	0	24610.1	0	0	22275.1	17304.6	20279.9
	(.)	(553.2)	(.)	(.)	(2029.5)	(1005.6)	(541.7)
Newton-Wellesley	8682.5	3677.5	9894.2	0	7754.0	10388.0	7505.2
	(433.6)	(225.4)	(1249.9)	(.)	(1308.3)	(287.2)	(203.0)
North Adams Regional	0	2543.9	0	0	0	0	4013.2
	(.)	(501.8)	(.)	(.)	(.)	(.)	(663.5)
Quincy MC	0	1551.2	0	0	0	5672.4	3597.0
	(.)	(553.2)	(.)	(.)	(.)	(799.6)	(549.3)
Steward Saint Anne's	0	1241.2	0	0	6848.0	8204.4	4860.8
	(.)	(491.7)	(.)	(.)	(1260.2)	(858.5)	(451.9)
South Shore	8577.8	3135.5	7649.7	0	7374.2	10134.7	6404.5
	(724.2)	(196.8)	(1332.4)	(.)	(1084.8)	(304.8)	(201.1)
Steward St. Elizabeth's MC	0	4354.7	0	0	8885.8	10427.9	8241.3
	(.)	(417.1)	(.)	(.)	(1362.5)	(408.9)	(328.4)
Saint Vincent	4973.1	2871.1	0	5172.7	5929.9	9631.6	5396.6
	(668.0)	(358.6)	(.)	(240.4)	(1291.7)	(553.3)	(217.9)
Sturdy Memorial	0	5628.2	0	0	5156.2	7138.8	5864.2
	(.)	(257.7)	(.)	(.)	(1507.5)	(805.6)	(318.9)
Clinton	0	0	0	0	0	0	3927.5
	(.)	(.)	(.)	(.)	(.)	(.)	(954.1)
Marlborough	0	2120.3	0	3007.6	8731.3	8507.0	4458.5
	(.)	(515.4)	(.)	(1023.2)	(2029.5)	(1184.3)	(528.9)
Winchester	0	3010.9	6241.1	0	4903.9	10028.8	6278.8
	(.)	(249.9)	(1613.6)	(.)	(1592.1)	(347.0)	(253.3)
Wing Memorial	0	1710.9	0	0	0	0	4050.9

	(.)	(819.5)	(.)	(.)	(.)	(.)	(785.7)
North Shore MC	10410.2	2964.4	0	0	7242.5	10180.5	5577.3
	(499.1)	(241.9)	(.)	(.)	(1413.2)	(766.5)	(276.4)
Boston MC	0	2828.3	0	0	5593.4	7310.2	4384.8
	(.)	(217.7)	(.)	(.)	(703.9)	(511.8)	(236.3)
Cambridge Health Alliance	0	1471.7	0	0	3048.7	5714.7	2274.3
	(.)	(255.3)	(.)	(.)	(1862.4)	(746.6)	(331.7)
MetroWest MC	0	1924.7	0	1568.5	5374.7	0	2393.0
	(.)	(250.3)	(.)	(905.2)	(1482.1)	(.)	(323.6)
Hallmark Health	0	2493.4	0	0	5204.6	9299.1	5339.6
	(.)	(591.9)	(.)	(.)	(1446.4)	(932.5)	(505.7)
Northeast	0	3438.0	0	3569.3	5297.2	9438.8	5466.6
	(.)	(199.7)	(.)	(877.1)	(1372.2)	(360.3)	(220.1)
Southcoast	7766.9	2317.5	7651.3	3144.4	8808.1	7674.9	4748.2
	(607.5)	(175.7)	(1493.9)	(995.9)	(848.7)	(381.3)	(194.5)
UMass Memorial MC	0	1100.5	11978.5	4698.0	7385.2	12213.5	6551.8
	(.)	(567.8)	(1363.8)	(256.9)	(820.0)	(491.5)	(223.7)
Berkshire MC	11209.2	0	0	0	9862.0	11012.2	10150.3
	(781.1)	(.)	(.)	(.)	(1136.8)	(881.1)	(460.7)
Lahey Clinic	0	4419.4	0	7065.2	9412.1	11588.1	7275.7
	(.)	(213.2)	(.)	(1085.3)	(813.8)	(415.4)	(227.6)
Mercy MC	0	1787.4	0	3485.0	0	6019.0	2965.1
	(.)	(310.4)	(.)	(792.6)	(.)	(914.4)	(374.5)
Steward Good Samaritan MC	0	2348.3	0	0	6416.5	7031.2	4392.8
	(.)	(249.0)	(.)	(.)	(1362.5)	(422.1)	(268.1)
Beth Israel Deaconess MC	0	3317.7	9080.4	0	7741.2	14187.4	8010.5
	(.)	(168.9)	(1007.2)	(.)	(755.4)	(261.5)	(171.6)
Observations	1267	14563	644	2188	3650	7894	31719

Standard errors in parentheses

C.4 OLS Robustness

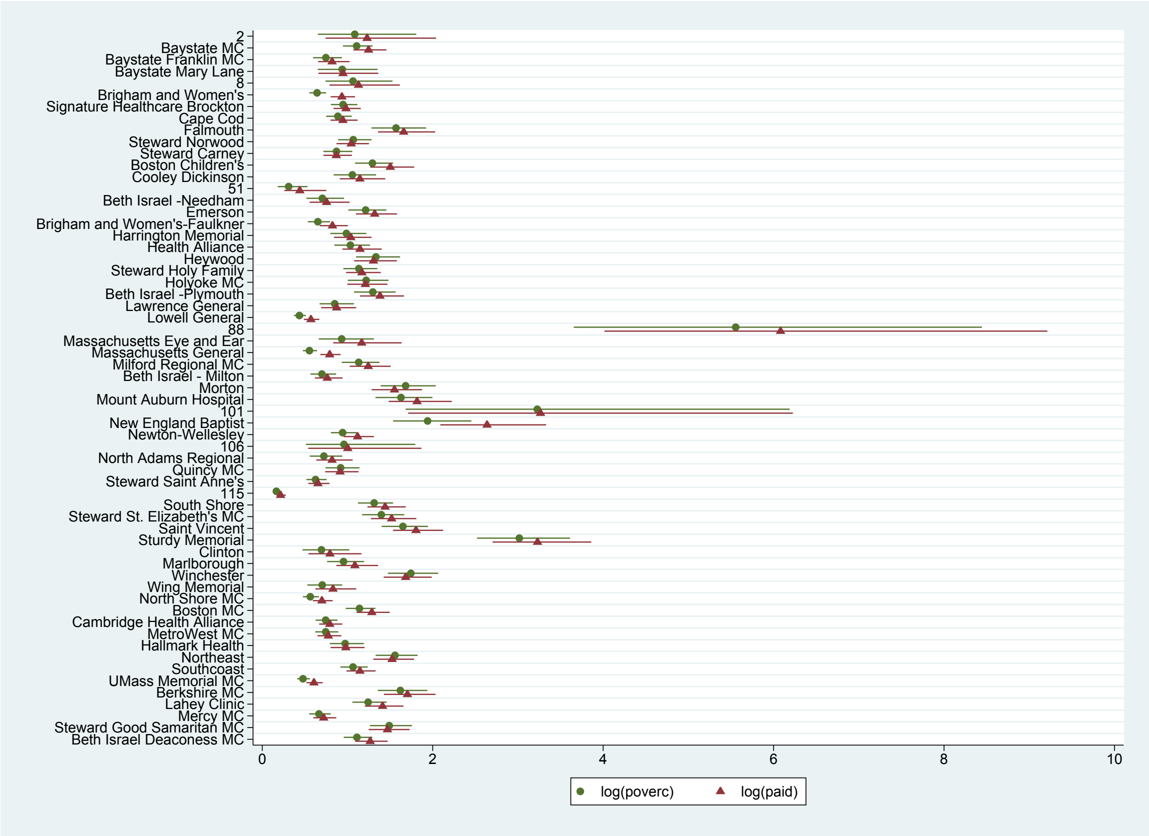


Fig. C.1. Restricted and unrestricted exponentiated coefficient plot for estimates of coefficients on hospital indicators. The green dots show the exponentiated coefficients of the restricted model, which regresses the log paid to charged ratio against hospital and insurer indicators (uninteracted). The red triangles show the exponentiated coefficients of the unrestricted model, which regresses log payments against log charges, hospital indicators, and insurer indicators (uninteracted). As the omitted hospital is Anna Jacques, the value of the exponentiated coefficient should be interpreted as the ratio of payments to charges relative to Anna Jacques. E.g.: a hospital with a coefficient equal to 0.7 has an average paid to charged ratio that is 70% of the same ratio at Anna Jacques.

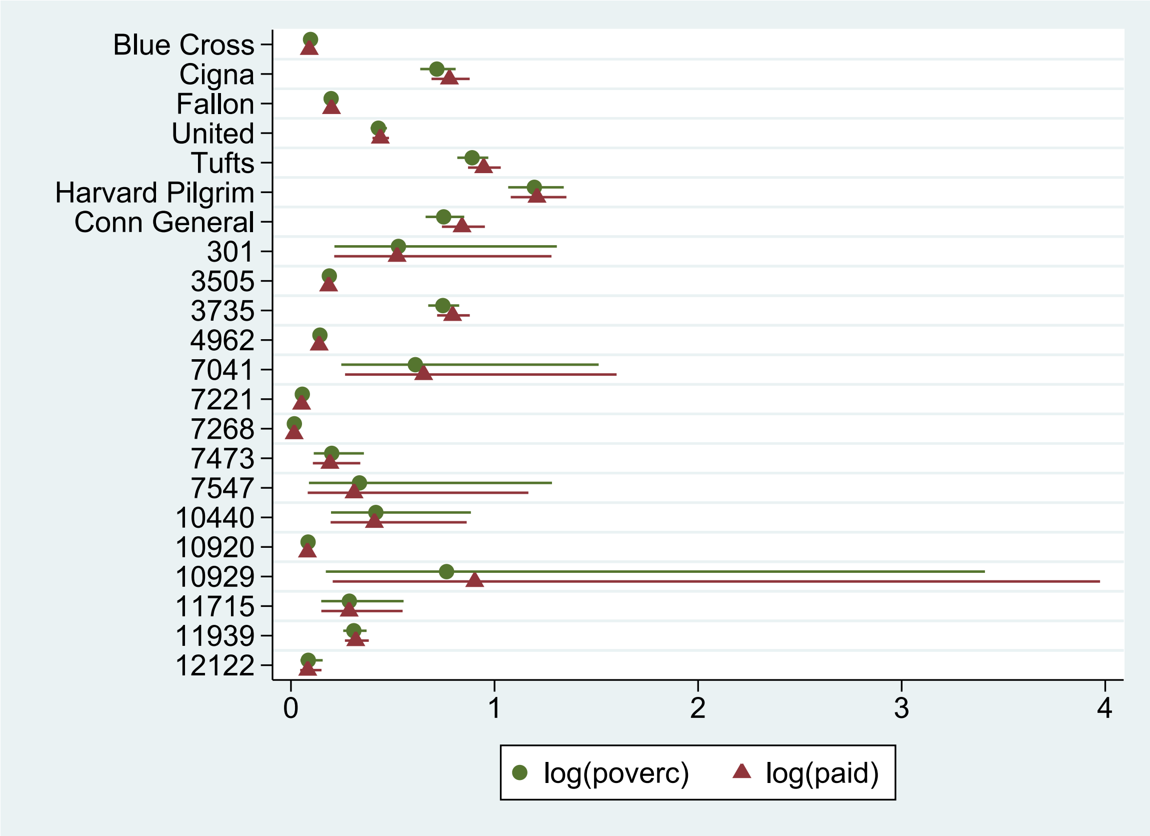


Fig. C.2. Restricted and unrestricted exponentiated coefficient plot for estimates of coefficients on insurer indicators. The green dots show the exponentiated coefficients of the restricted model, which regresses the log paid to charged ratio against hospital and insurer indicators (uninteracted). The red triangles show the exponentiated coefficients of the unrestricted model, which regresses log payments against log charges, hospital indicators, and insurer indicators (uninteracted). As the omitted insurer is Aetna, the value of the exponentiated coefficient should be interpreted as the ratio of payments to charges relative to those paid by Aetna. E.g.: an insurer with a coefficient equal to 0.7 has an average paid to charged ratio that is 70% of the same ratio for Aetna.

VITA

VITA

James' research interests are predominantly in the fields of Behavioral Economics and Econometrics. His research to date uses economic experiments to investigate the role of behavioral considerations in games. He complements the traditional experimental approach of analyzing treatment effects with structural econometrics. These techniques aim to uncover the unobservable motivations behind individuals decisions that economic theory predicts may be important. His work in econometrics for experiments proposes flexible methods for accounting for unobserved heterogeneity among participants that can be applied to already existing datasets. He has published in The European Economic Review, and presented at the Australia New Zealand Workshop on Experimental Economics, The ExperiMetrix workshop, and the Southern Economic Association annual meetings.

RESEARCH INTERESTS

Microeconomics, Game Theory, Industrial Organization, Behavioral Economics, Experimental Economics, Bayesian Econometrics, Directed Search

EDUCATION

In progress	Economics Ph.D., Purdue University, Krannert School of Management Adviser: Tim Cason
2011	Master of Economics - (Honours - H1), The University of Melbourne Adviser: Nikos Nikiforakis
2005-2010	Bachelor of Commerce, Degree with Honours, The University of Melbourne Adviser: Nikos Nikiforakis, Major: Economics
2005-2009	Bachelor of Engineering - (Honours - H1), The University of Melbourne Major: Chemical Engineering

HONORS AND AWARDS

2015	Robert W. Johnson Award for Distinguished Research Proposal, Purdue University *The Roommate's Dilemma: An experimental analysis of choice bracketing in games*
2011	CW Studentship, The University of Melbourne
2011	Kinsman Studentship, The University of Melbourne
2010	Victorian Department of Primary Industries Economics Honours Scholarship The University of Melbourne

PUBLICATIONS

Coordination with third-party externalities (with Nikos Nikiforakis). The European Economic Review, Volume 80, November 2015, pp 1-15

A Detailed Chemical Kinetic Model for Pyrolysis of the Lignin Model Compound Chroman (with Gabriel da Silva). AIMS Environmental Science, Volume 1, 2013, pp 12-25

Dimensional Cost Exponents for Novel Processing Equipment (with Alexandra Kingsbury, and Andreas Mönch), AusIMM (2012), Proceedings Project Evaluation 2012 , pp 189-194

WORKING PAPERS

How many games are we playing? An experimental analysis of choice bracketing in games

Monotonicity, Non-Participation, and Directed Search Equilibria (with Simon Loertscher)

WORK IN PROGRESS

Hospital-Insurer Bargaining Power and Negotiated Rates (with Amanda Cook)

Mixture models of behavior and nuisance parameters: a semi-parametric Bayesian approach (with Justin Tobias)

CSGTAP: Endogenous frictional unemployment in GTAP (with Thomas Hertel)

CONFERENCES, WORKSHOPS, AND SEMINARS

How many games are we playing? An experimental analysis of choice bracketing in games

- Annual Meeting of the Midwest Economic Association (April 2016)
- 85th Annual Meetings of the Southern Economics Association (November 2015)
- 2015 Spring School in Behavioral Economics, Rady School of Management, UC San Diego (poster session)

Mixture models of behavior and nuisance parameters: a semi-parametric Bayesian approach (with Justin Tobias)

- The ExperiMetrix workshop, University of Alicante (October 2015)

Hospital-Insurer Bargaining Power and Negotiated Rates

- Purdue University, Interdisciplinary Health Seminar
- 8th Annual Midwest Graduate Student Summit on Applied Economics, Regional and Urban Studies (2015)

Coordination games with third-party externalities (with Nikos Nikiforakis)

- Economic Theory and Experiments seminar series (2012). The University of Melbourne.
- The 6th Annual Australia New Zealand Workshop on Experimental Economics (2011). Monash University, Clayton
- National Honours Colloquium (2010). Australian School of Business, The University of New South Wales.

2014 IFREE Graduate Student Workshop in Experimental Economics – Attended

REFEREEING

The European Economic Review, Journal of the Economic Science Association

GRANTS

2011	Student Research Grant: *An Experimental Analysis of Equilibrium Selection in Coordination Games with Externalities*
2010	Student Research Grant: *An Experimental Analysis of Equilibrium Selection in Coordination Games with Externalities*

TEACHING

Purdue University

2015	Econometrics II (ECON 590), Summer semester – Teaching assistant Award: *Krannert Certificate for Distinguished Teaching*
2014	Behavioral Economics (ECON 471), Summer semester – Lecturer Award: *Krannert Certificate for Distinguished Teaching*
2013	Microeconomics (ECON 251), Fall Semester - Instructor for weekly review sessions
2012	Principles of Economics (ECON 210), Fall semester – Recitation instructor Award: *Krannert Certificate for Outstanding Recitation Teaching*

The University of Melbourne

2012	Microeconomics (Masters level) – Subject tutor Advanced Econometric Techniques (Masters level) – Subject tutor
2010	Chemical Engineering Management (Undergraduate level) – Subject tutor

AFFILIATIONS

American Economic Association, Southern Economic Association, Midwest Economics Association