Summer 8-1-2017

Three Essays on Experimental Economics and Individual Decision Making Under Risk

Xiaoxue Gao

Summer 8-1-2017

THREE ESSAYS ON EXPERIMENTAL ECONOMICS AND INDIVIDUAL DECISION MAKING UNDER RISK

Xiaoxue Gao

James C. Cox

Vjollca Sadiraj

Glenn W. Harrison

Rusty Tchernis

See next page for additional authors

Author
Xiaoxue Gao, James C. Cox, Vjollca Sadiraj, Glenn W. Harrison, Rusty Tchernis, Ira L. Leeds, Timothy M. Pawlik, Kurt E. Schnier, and John F. Sweeney

ABSTRACT

THREE ESSAYS ON EXPERIMENTAL ECONOMICS
AND INDIVIDUAL DECISION MAKING UNDER RISK

BY

XIAOXUE (SHERRY) GAO

AUGUST 2017

Committee Chair: Dr. James C. Cox

Major Department: Economics

The dissertation looks at three topics that involve experimental economics methods or individual decision making under risk: how do people make educational decisions when facing the risk of drop out; which models best characterize individuals' decision processes under risk; how can physicians improve discharge decisions to reduce the risk of unplanned readmissions.

In the first chapter, I introduce the risk of dropout into Spence's job market signaling model and test the modified model in the laboratory. I look at equilibria in the labor market and discuss the refinement based on the Cho-Kreps Intuitive Criterion. I derive the condition under which a separating equilibrium is the only perfect Bayesian equilibrium that survives the refinement and discuss the effects of workers' risk preferences on these equilibrium predictions. The data from lab experiments show that the market reaches the separating equilibrium more often when it is the only intuitive equilibrium. I also observe that, when the share of the low-ability type in the worker population decreases, or the cost to pursue a degree increases, the size of the wage premium for having the degree generally decreases. In the experiments, I use binary lottery tasks to elicit subjects' risk preferences to explain their strategies in the signaling games, and the analyses partially

confirm the prediction that more risk-averse individuals pursue a higher degree less frequently in the presence of dropout risks.

In the second chapter, as part of a joint project with Dr. Glenn W. Harrison and Dr. Rusty Tchernis, we apply the Bayesian econometric method to estimation of individual preferences under risk. We estimate a mixture model of Expected Utility Theory and Cumulative Prospect Theory using both simulated and observed binary lottery choices. We develop Markov Chain Monte Carlo algorithms to sample from the posterior distribution of parameters in the mixture model and compare the performances of different algorithms. The algorithms generally recover the true parameters used in the simulation, although some algorithms outperformed others in terms of efficiency. We also apply the algorithms to estimation using actual choice data. We find that 56.5% of the subjects can be characterized as consistent with Expected Utility Theory and 43.5% with Cumulative Prospect Theory. We find modest risk aversion among Expected Utility maximizers, and overweighting on the probabilities of extreme outcomes with very mild loss aversion among Cumulative Prospective Utility maximizers.

In the third chapter, coauthored with Dr. Ira L. Leeds, Dr. Vjollca Sadiraj, Dr. James C. Cox, Dr. Timothy M. Pawlik, Dr. Kurt E. Schnier and Dr. John F. Sweeney, we sought to define the association between information used for hospital discharge and patients' subsequent risk of unplanned readmission. De-identified data for patients from a tertiary academic medical center's surgical services were analyzed using a time-to-event model to identify criteria that statistically explained the timing of discharges. The data were subsequently used to develop a prediction model of unplanned hospital readmissions. Comparison of discharge behaviors versus the predictive readmission model suggested important discordance with certain clinical measures not being accounted for to optimize discharges. We suggest that decision-support tools for discharge may utilize variables that are not routinely considered by healthcare providers.

THREE ESSAYS ON EXPERIMENTAL ECONOMICS

AND INDIVIDUAL DECISION MAKING UNDER RISK

BY

XIAOXUE (SHERRY) GAO

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree
of
Doctor of Philosophy
in the
Andrew Young School of Policy Studies
of
Georgia State University

GEORGIA STATE UNIVERSITY

2017

ACCEPTANCE

This dissertation was prepared under the direction of Xiaoxue Gao's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair:	Dr. James C. Cox
Committee:	Dr. Glenn W. Harrison Dr. Vjollca Sadiraj Dr. Rusty Tchernis

Electronic Version Approved:

Sally Wallace, Dean
Andrew Young School of Policy Studies
Georgia State University
August, 2017

ACKNOWLEDGMENTS

I would like to thank my committee members for the constant advisements, heart-warming encouragement and tremendous support in writing this dissertation and in my time as a Ph.D. student at Georgia State University. I am truly blessed to have Dr. Cox, a rigorous scholar and a great story-teller, as my dissertation committee chair. Dr. Cox has guided me through every step in developing my first ever research project in experimental economics with great patience, and he has generously and unreservedly lent his support in all aspects of my Ph.D. life whenever I needed it. I was drawn to the study of individual decision making under risk and uncertainty by Dr. Harrison's lively lectures and enthusiastic research attitudes. From then on, Dr. Harrison's encouragement and advise have extended beyond classes and research, and have been a nurturing force in my growth as a scholar. In the many years working as her research and teaching assistant, Dr. Sadiraj has been the most endearing mentor a student could ask for. Dr. Sadiraj always had faith in me even when I doubted myself, and besides advising me on my research, helped me through so many difficulties with her caring, calming and wise words. As my advisor and the director of the Ph.D. program, Dr. Tchernis has always held me to high standards and helped me to be a better researcher as well as a better person. Together they shaped me into the scholar and person that I am now, by guiding me and being the best role models that I could hope for.

I would also like to thank Dr. Michael Price, Dr. Garth Heutel, Dr. Spencer Banzhaf, Dr. Yongsheng Xu and Dr. Daniel Kreisman for advising me while I was on the job market, Kevin Ackaramongkolrotn, Lucy Gentles, Bess Blyer and Dr. Todd Swarthout for their kind help in running my experiments and in my job search, and Dr. Lisa Armistead, Dori Neptune and Katrina Helz for providing fellowship support for my last year in the program. I am also truly appreciative of the support and encouragement from my colleagues and friends: Alexander, Anomitro, Bo, Dong, Guanlin, Lily, Prithvijit, Rhita, Valentina, Xi, Xiaoran, Xiaoxi, Xilin, Zhixuan. I also want to thank Noodle, the golden retriever of Dong and Zhixuan, for always lightening me up with his unreserved affections.

At last, I would like to thank my family with the deepest love. To my loving and caring mother, Hong Liu, I could not have gone half as far without your love and support. To my wise and humorous father, Zhonghai Gao, thank you for always making everything less scary with your jokes and long talks with me. To my aunts Chunli Liu, Jing Gao, Ping Gao, Minlan Huang and my uncles Yong Meng, Zhengqui Li, thank you all for supporting me and doting me like your own child. To my grand parents Naicui Lv, Futian Gao and Qingxian Liu, thank you for always spoiling me with so much food and love whenever I visited home. Finally, I want to dedicate this dissertation to my dearest grandmother who passed away three years ago, Qinglian Li. I really miss you and truly regret not being there for you when you were battling with illness. I consider myself the luckiest child growing up with your care and love.

Contents

[1] Coauthored with Dr. Glenn W. Harrison and Dr. Rusty Tchernis. I thank Dr. Glenn W. Harrison and Dr. Todd Swarthout for providing the observed data used in Section 2.5.

[2]Coauthored with Ira L. Leeds , Vjollca Sadiraj, James C. Cox, Timothy M. Pawlik, Kurt E. Schnier and John F. Sweeney, published in The American Journal of Surgery 213 (2017) 112-119. Permission to reproduce this paper is granted by ©2016 Elsevier Inc. The published paper can be found at http://dx.doi.org/10.1016/j.amjsurg.2016.03.010.

List of Tables

List of Figures

Chapter 1

Job Market Signaling: An Experimental Study of Education Degree as an Imperfect Signal

1.1 Introduction and Literature Review

In his paper “Job Market Signaling”(Spence (1973)), Spence looks at how education can act as a signal of innate ability when workers with higher abilities self-select to pursue more education. As an important measure of education level, when a higher degree is perceived as a signal of greater ability, it can result in a degree wage premium– the increase in earnings from having a higher degree between otherwise similar workers.[1] Spence proposes a necessary condition for the self-selection process and the existence of the degree wage premium: the inverse relationship between education costs and abilities. However, when we look at degree programs, it is not uncommon to see students fail to meet the degree requirements and drop out of the program involuntarily – a risk of the education investment that could also trigger the self-selection process, even when the aforementioned necessary condition is violated.

This paper extends the signaling theory by introducing the dropout risk into the model and explores how it can trigger the self-selection process among workers and generate a degree wage premium when, ceteris paribus, a worker with lower ability is at higher risk of dropping out.

[1] There has been persistent evidence of the wage premiums associated with educational degrees in US labor markets. Early evidence includes Hungerford and Solon (1987), Belman and Heywood (1991), Heywood (1994), Hungerford and Solon (1987) and Jaeger and Page (1996). Using 1990 CPS data, Park (1999) finds that, among the individuals who have the same years of schooling, there are significant earning gains of 9, 11 and 21% from obtaining a high-school diploma, an associate’s degree, and a bachelor’s degree, respectively.

Specifically, I define the cost as tuition and forgone earnings in pursuit of a degree, and the dropout risk as the likelihood of a worker not earning the degree after enduring the costs.[2] In making this differentiation, I refrain from the risk neutral assumption and discuss how the agents' risk attitudes affect the equilibrium in the market. In fact, with education cost defined to include the forgone earnings, there might not be much of a difference in the costs between workers with different abilities;[3] thus, the inverse relationship between ability and dropout risk may be indispensable in triggering the self-selection process among workers and in explaining the observed degree wage premium.

In this paper, I first formalize the idea in a theoretical framework and then test the predictions with laboratory experiments. In the theoretical section, I look at a simplified scenario in which there are two types of workers in the labor market – those with high and those with low abilities. The workers need to decide whether to pursue a certain degree at some cost; to focus on the signaling role of the degree, I assume that pursuing the degree does not improve ability or productivity. After incurring the cost, there is a risk that workers fail the degree requirements and not earn the degree, and the risk is greater for workers of the low-ability type. I find that even when the costs are the same for both types, which violates the original Spence condition, in the presence of heterogeneous dropout risks, a separating equilibrium[4] still exists and may be the only Perfect Bayesian Nash Equilibrium that satisfies refinement based on the Cho-Kreps Intuitive Criterion[5]. The theoretical discussion predicts that, the more risk-averse a worker is, the less likely he or she

[2] I assume the *ex-ante* probability of dropout is known to both workers and employers here. An interesting extension is to consider the situation in which the probability is unknown, and look at the effects of uncertainties instead of risks of dropout.

[3] Arguably, the less able can spend more years in the degree program and, therefore, endure greater costs; however, most programs post an expiration date on the credits earned by the students, so there is a limit to how much the costs can increase. Another argument is that outstanding students might be able to acquire scholarships and lower their cost, but the earnings they have forgone may be greater too.

[4] In a separating equlibrium, the high type pursues the degree, while the low type does not; recognizing this, employers offer higher wages to degree holders in the labor market. In a pooling equlibrium, neither the high type nor the low type pursues the degree, and employers offer the same wages to them since they cannot distinguish one type from the other.

[5] Cho and Kreps (1987) came up with the "Intuitive Criterion" to rule out the equilibrium supported by unreasonable out-of-equilibrium beliefs – in the current context, the pooling equilibrium. Details of this criterion will be discussed in Section 1.2.

will pursue the degree in the presence of the dropout risk; while moderate risk aversion or loving affects only the degree of separation in the market, extreme risk attitudes could potentially result in the non-existence of a separating equilibrium. Data from the lab experiments generally support the refinement and the comparative statics of the model and also suggest that subjects learn to play the game throughout the many market periods. In the experiments, I use binary lottery tasks to elicit subjects' risk preferences to account for their strategies in the signaling games, and the analyses partly confirm the prediction that more risk-averse individuals pursue a higher degree less frequently.

Although developed in the context of job market signaling, this paper also adds to the literature on general signaling games by introducing *type-dependent* stochastic noise. Matthews and Mirman (1983) were the first to look at stochastic signaling games by introducing demand shocks to the incumbent's price choice in the entry limit pricing game.[6] Landeras and Villarreal (2005) present a screening model with performance noise to the educational credentials and conclude that over-education is made worse by the noise in students' credentials. In more recent work, de Haan et al. (2011) and Jeitschko and Normann (2012) introduce noises to general signaling games with different focuses: the former on the effects of different noise levels and the latter on prior distributions of sender type. However, these studies commonly assume independence between the distribution of noise and sender type, so they still rely on the difference in signaling costs to induce signaling behavior. In contrast, the current work introduces dependence between noise distribution and ability. In this setup, the type of senders who are more desirable to the receivers have greater control over the signal generating process than the less desirable type, which can induce signaling behavior in the market even when the signaling costs are the same.[7]

[6]Based on their work, Carlsson and Dasgupta (1997) propose the "noise-proof" criterion as the equilibrium selection tool in deterministic signaling games.

[7]Another paper that introduces such dependence is Regev (2012), which introduces a test that workers can take regardless of their educational backgrounds. However, due to their assumptions that the test costs workers nothing and that the likelihood of passing cannot be improved by education, he arrives at the completely different prediction that a separating equilibrium does not exist and that workers do not self-select into different education levels. As I will argue in Section 1.2, the existence of a signaling cost is essential in inducing a separating equilibrium.

The current paper also tests the theoretical model in the laboratory, which adds to the experimental literature on testing equilibrium prediction and selection in signaling experiments. Since the 1980s, *deterministic signaling models* have been tested using lab experiments in both market and non-market environments. Examples of non-market experiments include Brandts and Holt (1992), Banks et al. (1994) and Potters and van Winden (1996). Among market experiments, Cadsby et al. (1990, 1998) examine signaling in corporate finance models in which firms can signal their type with investment decisions; Cooper et al. (1997) and Cooper and Kagel (2003) look at entry-limit pricing games in which the incumbent can signal its cost type or the market profitability to potential entrants with prices; and Posey and Yavas (2007) study health insurance screening, while Miller and Plott (1985) and Kübler et al. (2008) look at generic product markets with information asymmetry. The commonly asked question is: when the parameters are within the experimenter's control, do subjects behave as the theoretical equilibrium predicts? The above-mentioned studies find that subjects behave generally consistently with the Bayesian Nash equilibrium predictions; however, the extent to which observations are consistent with theoretical predictions is never 100%. They also find that subjects sometimes, but not always, play the more refined equilibrium when the game has multiple equilibria, depending on the complexity of games and the design of experiments. [8]

Jeitschko and Normann (2012) and de Haan et al. (2011) are the first to test *stochastic signaling models* in a lab environments. In de Haan et al. (2011), the authors fix the population distribution while changing the variance of the noise at low, intermediate and high levels. They find that the noise level systematically affects the signal behavior of high-quality sellers. Jeitschko and Normann (2012) compare subjects' behavior in deterministic versus stochastic settings when the population distribution is either 33% or 67% the high type, keeping the noise distribution the same in the stochastic treatments. Their data support the comparative statics well, but generally do not converge completely to the theoretical equilibrium predictions. In both experiments, dif-

[8] The intuitive criterion is a popular refinement measure tested in these experiments. Brandts and Holt (1992, 1993) and Cooper and Kagel (2003) find evidence indicating subject behavior may contradict the Intuitive Criterion and follow history-dependent learning processes when the payoff structures make it harder for them to apply this forward-induction-based reasoning.

ferent types have different signal costs but share the same distribution over the signal noise – the same feature that differentiates their studies from the current paper in terms of the theoretical development of signaling models. In addition, the experiments in these two papers do not consider subjects' risk preferences when examining their decisions in the signaling games, despite of the stochastic nature of the games. In the current experiments, I include a second stage in which subjects' risk attitudes are elicited using binary lottery choices to explain their play of the signaling games in the first stage.[9]

In Section 1.2, I develop the model to accommodate both heterogeneous costs and heterogeneous noises between the two types, and I discuss the implications of homogeneous costs as a special case. In Section 1.3, I describe the experiment design and procedures;[10] to focus sharply on the role of dropout risk in signaling behavior, I have kept signal costs the same while imposing different dropouts risks on both types in the experiments. Section 1.4 presents the data and results, and Section 1.5 concludes.

1.2 Model

Suppose that in a labor market, we have two types of workers, either high-ability or low-ability, and the proportion of the low ability type is μ. Assume that different abilities lead to different productive efficiencies when workers are hired: the more able workers will have higher productivity θ_h, while the less able workers will have lower productivity θ_l, and we have $\theta_h > \theta_l$. Workers can choose to pursue an education degree at a cost of c; assume that higher ability leads to a lower cost of the degree, so we have $0 < c_h < c_l$.

If a worker decides to incur the cost and pursue the degree, with probability λ, she will fail to meet the degree requirements and drop out; assume that higher ability also leads to a lower risk of dropout, so we have $0 \leq \lambda_h < \lambda_l$. Denote the high type's choice as $e_h \in \{0,1\}$ and the low

[9] To my best knowledge, in the line of research on stochastic signaling games, this paper is the first to explicitly tests the effects of risk attitudes on agents' play of such games.

[10] The experiment procedures of the current paper are similar to those in Kübler et al. (2008); their experiments can be seen as a parametrization of the current model with different education costs and no dropout risk.

type's as $e_l \in \{0,1\}$, where 1 is pursuing and 0 is not pursuing. Denote the education outcome as $e_o \in \{0,1\}$, where 1 means a worker has the degree and 0 means not; then, we have

$$\lambda_h = \text{Prob}\{e_o = 0 | e_h = 1\}, \quad \lambda_l = \text{Prob}\{e_o = 0 | e_l = 1\}.$$

Assume that employers are competing to hire workers – if they can directly observe a worker's type at the time of hiring, they will offer a wage that is equal to her productivity.[11] However, ability is not directly observable at the time of hiring, so employers will offer the worker a wage that they *believe* to be consistent with her productivity, based on some observable attribute – in this case, whether she has the degree. Assume that employers can observe a worker's education outcome but not her choice; therefore, they cannot *credibly* differentiate the dropouts from those who have decided not to pursue the degree. Once hired, a worker's type will eventually be revealed in production, and employers can update their beliefs for the next round of hiring.[12] Suppose that employers believe that a worker *without* the degree has a likelihood μ_0 of being the low type, and a worker *with* the degree has a likelihood μ_1 of being the low type; that is,

$$\mu_0 = \text{Prob}\{\theta = \theta_l | e_o = 0\}, \quad \mu_1 = \text{Prob}\{\theta = \theta_l | e_o = 1\}.$$

[11] If employers have more bargaining power and workers are competing for jobs, then the equilibrium wage will be driven down to the reservation wage regardless of a worker's productivity. The assumption that employers will offer wages equal to a worker's productivity comes from the neoclassical profit-maximization problem of competitive firms. However, this assumption is not crucial to the current model. We can think of θ_h and θ_l as how much a firm is willing to pay to hire the high type and the low type, and that firms are willing to pay more to hire a worker with higher ability.

[12] An alternative way to credibly reveal a worker's type is to strike a long-term contract with the same initial wages, and increase the wage of the high type based on performance later. However, under the assumption that employers are competing for workers, a high-type worker is unlikely to accept such a contract from one employer, when another is willing to offer what she deserves from the beginning of the contract.

1.2.1 Risk Neutral Agents

Assume that education does not improve workers' productivity[13] and that all agents are risk-neutral. The strategic decisions in this model can be characterized as follows:

i) Based on prevalent wage offers, a worker will maximize her expected payoff (defined as wage net of education cost), and a type i worker will choose to pursue the degree only when this leads to a higher expected payoff:

$$w(0) < (1-\lambda_i)\cdot(w(1)-c_i)+\lambda_i\cdot(w(0)-c_i),\ i\in\{h,l\}.$$

ii) Based on post-hiring verification, employers' beliefs μ_0 and μ_1 should be updated using Bayes' rule whenever possible.[14] Employers will offer a wage that is equal to a worker's expected productivity, given her degree status:

$$w(e_o)=\begin{cases}(1-\mu_0)\ \theta_h+\mu_0\ \theta_l, & e_o=0\\ \\ (1-\mu_1)\ \theta_h+\mu_1\ \theta_l, & e_o=1.\end{cases}$$

The market is in Perfect Bayesian Equilibrium if workers' educational decisions and employers' wage offers sustain each other. There are potentially three pure-strategy equilibria: both types choose not to pursue a degree ($e_h=e_l=0$); the high type pursues, while the low type does not ($e_h=1, e_l=0$); and both workers choose to pursue the degree ($e_h=e_l=1$). For the sake of brevity, I will refer to the three cases as "pooling not to pursue", "separating" and "pooling to pursue" hereafter. In Appendix C, I discuss the decision processes outlined in i) and ii) for each case. Here, I present the equilibrium predictions in the first proposition.

[13]The purpose of this assumption is to keep the discussion sharply focused on the signaling process, and it should not be taken as a complete denial in the productivity-improving function of education. Note that the framework in this section allows the incorporation of productivity-improving function, which will shift the wage offer $w(1)$ up by the increment in productivity.

[14]Beliefs associated with out-of-equilibrium strategies are not restricted by the Perfect Bayesian Equilibrium solution, as employers cannot frequently observe such behavior to verify the type, but refinement on out-of-equilibrium beliefs will be applied later.

Proposition 1 Pooling not to pursue is always an equilibrium of this market, while separating is an equilibrium iff the education costs and dropout risks satisfy

$$\frac{c_h}{1-\lambda_h} \leq \frac{\mu}{\mu+\lambda_h(1-\mu)}(\theta_h-\theta_l) \leq \frac{c_l}{1-\lambda_l}.$$

When the above condition is satisfied, the inefficient case in which both workers pursue the degree is *not* an equilibrium.

Although pooling not to pursue is always a Perfect Bayesian Equilibrium, it is supported by employers' out-of-equilibrium beliefs that associate a degree holder with the low type with some probability. Such beliefs are particularly counterintuitive when the low type's cost or dropout risk is so high that, in expectation, they would not benefit from pursuing the degree, even when the employers offer the highest possible wage, θ_h, to degree holders. To refine multiple equiliria in signaling games, Cho and Kreps (1987) came up with the "Intuitive Criterion" to rule out the equilibrium supported by unintuitive out-of-equilibrium beliefs. In the current context, the criterion requires that employers do not associate degree holders with the low type, when it is not profitable for the low type to deviate and pursue the degree, while it might be for the high type. In Appendix C, I derive the conditions under which we can refine the pooling equilibrium based on the Intuitive Criterion, and the following proposition presents my conclusion.

Proposition 2 Pooling not to pursue is *not* an intuitive Perfect Bayesian Equilibrium if the costs and dropout risks satisfy

$$\frac{c_h}{1-\lambda_h} < \mu(\theta_h-\theta_l) < \frac{c_l}{1-\lambda_l}.$$

Combined with the condition for a separating equilibrium to exist, separating will be the *only* intuitive Perfect Bayesian Equilibrium when

$$\frac{c_h}{1-\lambda_h} < \mu(\theta_h-\theta_l) < \frac{\mu}{\mu+\lambda_h(1-\mu)}(\theta_h-\theta_l) < \frac{c_l}{1-\lambda_l}.$$

Fixing θ_h, θ_l and μ, two interesting cases arise if we change the assumptions on education cost and dropout risk. If we eliminate dropout risk from the model such that $\lambda_h = \lambda_l = 0$, the above condition adapts to

$$c_h < \mu(\theta_h - \theta_l) < \theta_h - \theta_l \leq c_l.$$

In contrast to the original condition (other parameters constant), without dropout risks, the cost for the low types c_l will have to be higher to deter them from pursuing the degree, while the cost for the high types will not need to be that low to encourage them to pursue the degree.

In the second case, if we eliminate the difference in education costs between the two types such that $c_h = c_l = c$, then the condition adapts to

$$1 - \lambda_h > \frac{1}{\mu} \cdot \frac{c}{\theta_h - \theta_l} > \frac{\mu + \lambda_h(1-\mu)}{\mu} \cdot \frac{c}{\theta_h - \theta_l} \geq 1 - \lambda_l.$$

That is, even when the education cost is not negatively correlated with workers' abilities, as long as the dropout risk is sufficiently low for the high type and sufficiently high for the low type, workers can be separated by their educational choices, and, consequently, employers can use the degree as a signal of high ability. Note that education costs should always be positive (non zero); otherwise, the low type will always pursue the degree, even when their dropout risk is very high.

Corollary 1 When education costs are not zero and are the same for both types of workers, separating is the only intuitive Perfect Bayesian Equilibrium when the dropout risk is sufficiently low for the high type and sufficiently high for the low type:

$$1 - \lambda_h > \frac{1}{\mu} \cdot \frac{c}{\theta_h - \theta_l} > \frac{\mu + \lambda_h(1-\mu)}{\mu} \cdot \frac{c}{\theta_h - \theta_l} \geq 1 - \lambda_l.$$

1.2.2 Effects of Risk Attitude

The discussion up to this point has assumed that agents in the labor market are risk-neutral; however, the element of dropout risk in pursuit of an education degree makes the discussion of workers' risk attitudes quite relevant.[15] Assume that w and c are perfect substitutes in workers' utility function $u(w,c) = u(w-c)$. Also, assume that workers' decision can be characterized by expected utility theory.[16] Then, in the separating equilibrium, workers' incentive constraints in terms of expected utility are:

$$(1-\lambda_h)\cdot\mathbf{u}\big(w(1)-c_h\big)+\lambda_h\cdot\mathbf{u}\big(w(0)-c_h\big)\geq\mathbf{u}\big(w(0)\big) \qquad \text{High Type IC}$$

$$(1-\lambda_l)\cdot\mathbf{u}\big(w(1)-c_l\big)+\lambda_l\cdot\mathbf{u}\big(w(0)-c_l\big)\leq\mathbf{u}\big(w(0)\big) \qquad \text{Low Type IC}$$

Suppose that the workers are risk-averse $\mathbf{u}'(\ )>0$ and $\mathbf{u}''(\ )<0$; then, for $i=h,l$ we have

$$\begin{aligned}(1-\lambda_i)\mathbf{u}\big(w(1)-c_i\big)+\lambda_i\mathbf{u}\big(w(0)-c_i\big) &< \mathbf{u}\big((1-\lambda_i)(w(1)-c_i)+\lambda_i(w(0)-c_i)\big)\\ &= \mathbf{u}\big((1-\lambda_i)w(1)+\lambda_i w(0)-c_i\big)\end{aligned}$$

Comparing the two incentive constraints with their counterparts in the risk-neutral case, the IC for the low type is less stringent:

Lemma 3 If $(1-\lambda_l)w(1)+\lambda_l w(0)-c_l\leq w(0)$, then $(1-\lambda_l)\mathbf{u}\big(w(1)-c_l\big)+\lambda_l\mathbf{u}\big(w(0)-c_l\big)<\mathbf{u}\big((1-\lambda_l)w(1)+\lambda_h w(0)-c_l\big)\leq\mathbf{u}\big(w(0)\big)$

The lemma states that, when the low-type workers are risk-averse, the education cost or dropout risk does not need to be as high as in the risk-neutral case in order to deter them from pursuing the education degree and to induce separation. However, the IC for the high type is more stringent:

[15]Hiring a worker, in this model, is essentially buying a lottery with two outcomes: θ_h and θ_l; if we refrain from the risk-neutral employer assumption, the wage offer will be the certainty equivalents rather than the expected values of the lotteries, conditional on employers' beliefs. I will leave the effect of the employer's risk attitudes to be empirically evaluated in the experiment results section, while focus on the risk attitudes of workers in this section.

[16]Under Expected Utility Theory, an agent's risk attitude is solely characterized by the curvature of the utility function. Given the set up of this model, probability weighting can be incorporated into workers' incentive constraints by replacing probabilities with decision weights.

Lemma 4 If $(1-\lambda_h)\mathbf{u}\big(w(1)-c_h\big)+\lambda_h\mathbf{u}\big(w(0)-c_h\big)\geq\mathbf{u}\big(w(0)\big)$, then $\mathbf{u}\big((1-\lambda_h)w(1)+\lambda_h w(0)-c_h\big)>(1-\lambda_h)\mathbf{u}\big(w(1)-c_h\big)+\lambda_h\mathbf{u}\big(w(0)-c_h\big)\geq\mathbf{u}\big(w(0)\big)$; that is, $(1-\lambda_h)w(1)+\lambda_h w(0)-c_h>w(0)$.

Risk aversion among high type requires the education cost or the dropout risk to be lower than in the risk-neutral case, in order to encourage the high type to pursue the degree.[17] Therefore, risk aversion moves the boundaries of the parameter regions in which the separating equilibrium exists, while the overall structure of the analysis stays the same.

The discussion of workers' risk attitudes thus far rests on an implicit assumption that all workers have the same risk preferences. When workers have heterogeneous risk attitudes, the situation becomes complex and the separating equilibrium may cease to exist. Consider the situation in which education costs cannot be effectively differentiated between the high and low types – that is, $c_h=c_l=c$. Reorganize the existence condition of the separating equilibrium in terms of $\frac{c}{\theta_h-\theta_l}$ as follows:

$$\frac{\mu}{\mu+\lambda_h(1-\mu)}(1-\lambda_l)\leq\frac{c}{\theta_h-\theta_l}\leq\frac{\mu}{\mu+\lambda_h(1-\mu)}(1-\lambda_h).$$

If we assume the high-type workers are *risk-averse*, then the IC for the high type in the separating equilibrium will be more stringent, causing the right bound for $\frac{c}{\theta_h-\theta_l}$ to shift left. If we assume the low-type workers are *risk-loving*, then the IC for the low type will also be more stringent, causing the left bound to shift right. These two effects, if large enough, will lead to a contradiction between the two ICs in the separating equilibrium. In this situation, given c, θ_h, θ_l and μ, a separating equilibrium cannot be induced by the dropout risks, λ_h and λ_l.

In the situation discussed above, it takes both extreme risk loving among low productive workers and extreme risk aversion among high productive workers to make the existence of a separating equilibrium impossible. Laboratory research on subjects' risk attitudes, such as Holt and Laury (2002), Hey and Orme (1994) and Harrison and Rutström (2008), generally find small to modest risk aversion among subjects in the laboratory environment, with stakes similar to what will be

[17] Meanwhile, if we are looking at risk-loving workers, the conclusions will be reversed for these two types, respectively.

used in my experiment. Also, by the design of the experiment, I will be able to observe each subject's decisions as both high-type and low-type workers, so comparing signaling decisions made by the *same* subject between the two types will be equivalent to holding risk attitudes the same between types. To make comparisons across different subjects, risk attitudes will be elicited from each subject after they are done with the signaling game and will be used to explain between-subject differences in wage offers and education choices.

1.3 Experiment Design

The experiments are designed to test the model developed above and to answer the following questions: Will subjects learn to play the separating equilibrium in this labor market? Will the market reach the separating equilibrium more often when pooling not to pursue is unintuitive? Can risk attitudes account for subjects' decisions as workers and as employers? To focus sharply on the role that dropout risk plays in inducing separation, education costs are set to be the same for both types of workers.[18] Varying the education cost and population distribution of workers, while keeping workers' productivities and dropout risks constant, we have five treatments, shown in Table 1.1. In treatments 1, 2 and 3 at the top half of Table 1.1, education costs are kept constant, while the percentage of low types in the worker population decreases linearly by 25%. Therefore, I refer to them as *Population Treatments*. In treatments 4, 2 and 5 at the bottom half, the percentage of low types is kept constant, while education costs increase linearly by \$2. Therefore, I refer to them as *Cost Treatments*. Note that Treatment 2 is at the center of both treatment groups and will serve as the baseline for pairwise hypothesis tests later.

[18]In the experiments reported in Kübler et al. (2008), education costs are higher for low types and no dropout risk is imposed. Their experiments can be seen as a parametrization of the model developed in Section 1.2, in which $c_h > c_l > 0$ and $\lambda_h = \lambda_l = 0$. The procedures used in the current experiments are very similar to those in Kübler et al. (2008); therefore, the results in Kübler et al. (2008) can lend some insights in the discussion of the results from the current experiments.

Table 1.1: Parameterization in Each Treatment

Treatment	Low Type	Education Cost	Productivity		Dropout Risk		Is Equilibrium?	
			High Type	Low Type	High Type	Low Type	Separating	Pooling[1]
1	***75%***	$5	$25	$10	10%	90%	Yes	Yes but Unintuitive
2	***50%***	$5	$25	$10	10%	90%	Yes	Yes but Unintuitive
3	***25%***	$5	$25	$10	10%	90%	Yes	Yes
4	50%	***$3***	$25	$10	10%	90%	Yes	Yes but Unintuitive
2	50%	***$5***	$25	$10	10%	90%	Yes	Yes but Unintuitive
5	50%	***$7***	$25	$10	10%	90%	Yes	Yes

[1] Pooling to pursue is not an equilibrium in any of these treatments, so in the current table, as well as in the following discussion, pooling is referring to “pooling not to pursue”.

Table 1.2: Cho-Kreps Intuitive Refinement of Pooling Equilibrium

Treatment		**Population Treatments**			**Cost Treatments**		
		75%, $5	50%, $5	25%, $5	50%, $3	50%, $5	75%, $7
Pooling Wage		13.75	17.5	21.25	17.5	17.5	17.5
Expected Payoff if Deviate (Highest Possible[1])	**High Type**	18.875	19.25	19.625	21.25	19.25	17.25
	Low Type	9.875	13.25	16.625	15.25	13.25	11.25
Pooling at Not Pursuing		Unintitive	Unintuitve	Intuitive	Unintuitve	Unintuitve	Intuitive

[1] The highest possible expected payoff is calculated under the most favorable wage offer for degree holders; that is, employers offer $25 to workers who have a degree.

In all five treatments, pooling not to pursue and separating are both Perfect Bayesian Equilibria, while pooling to pursue is not; therefore, “pooling" refers to “pooling not to pursue” in the discussion hereafter. However, the five treatments differ in whether pooling is unintuitive; Table 1.2 applies the Intuitive Criterion to the pure-strategy pooling equilibrium of these treatments. The row labeled “Pooling Wage" shows the wage offer to workers without the degree from risk-neutral employers in the pooling equilibrium; the next two rows show the expected payoff of each type should employers offer $25, the highest possible wage, to workers who deviate from pooling and successfully earn the degree. In the first two treatments (columns) of each treatment group, pooling is unintuitive since deviation yields a lower expected payoff for the low type even under the most favorable out-of-equilibrium wage offer, while the high type can get a higher expected payoff. In the last treatment of each treatment group, neither type can profit from deviation in expectation, and pooling cannot be refined by the Cho-Kreps Intuitive Criterion.

Previous signaling experiments have shown that the separation between the high and low type's signaling decisions was never 100%. The purpose of these treatments is to see if changes in share of low types in the population or signal cost will make separation more or less complete. In addition, within each treatment group, although the changes in the treatment parameter are linear, pooling equilibrium is unintuitive in the first two treatments but intuitive in the last treatment. Therefore, with this design, the effectiveness of the Intuitive Criterion will manifest as a more dramatic increase in pooling (or, equivalently, a much less complete separation) when the cost changes from \$5 to \$7 than when it changes from \$3 to \$5 in the Cost Treatments, and when the share of low types in the population changes from 50% to 25% than when it changes from 75% to 50% in the Population Treatments.

Ten sessions were conducted between December 2015 and March 2017 in the lab at Georgia State University's Experimental Economics Center. A total of 240 subjects were recruited from among the undergraduate students at the university. Each session lasted approximately two hours. Twenty-four subjects participated in each session and two sessions were ran for each treatment. The market parameters were fixed within each session. To avoid potential confounding effects from contexts such as "education degree", "employers" or "workers", we used neutral wording in the experiments. Workers are referred to as sellers who are selling products with either high values or low values, and employers are referred to as buyers who are bidding for the products. Pursuing the degree is described as testing the product, and the degree itself is referred to as a quality certificate. The subject instructions for Treatment 2 are attached as Appendix D.

In each session, the subjects play the signaling game for 32 periods; the 32 periods are divided into 4 blocks $\times$ 8 periods per block, and each subject assumes the same role within a block but might switch roles between blocks. The block and role switch design is common in signaling experiments: it allows us to observe each subject's decisions as both buyer and seller; and it can help prompt subjects understanding of the game from different angels.[19] Table 1.3 shows a decomposition of the block design and role assignment rules within each session.

Table 1.3: Role Assignment in Each Session

	Periods			
Subject ID [1]	1 - 8	9 - 16	17 - 24	25 - 32
1 - 8	*Seller*	Buyer	Buyer	Seller
9 - 16	Buyer	*Seller*	Buyer	Buyer
17 - 24	Buyer	Buyer	*Seller*	*Buyer*

[1] Subject ID was randomly assigned, so subjects with adjacent IDs do not necessarily take adjacent seats.
[2] Subjects were merely informed that their roles would switch back and forth over the 32 periods, but were not told the specifics of the block design or the role switching rules.

When each period begins, each seller will get one product with either a high or a low quality. Then, two buyers will be randomly and anonymously matched with the seller to undergo a first-price sealed-bid auction on her product.[20] The qualities of the products that each subject gets to trade are assigned in the following way (using Treatment 2 as an example):

1. In each period, half of the eight sellers will get the high-quality products and the other half will get the low-quality products. This information is *included* in the instructions.

2. In the eight periods of each block, each seller will get high-quality products for four periods and low-quality products for the remaining four periods; each buyer will be matched with high quality sellers for 4 periods and with low quality sellers for the rest 4 periods. This information is *excluded* from the instructions, and the sequences of product qualities that a subject gets to trade do not follow any patterns. Since subjects do not know the specifics of the block and role switching design, the possibility that they will figure out the sequence of quality assignments over the 32 periods is minimal.

[19] As will be introduced later, one of the 32 periods will be chosen to determine their payoff. Under the assumption that agents are expected utility maximizers, the role switching and block design combined with this payoff mechanism should not create incentives for portfolio decisions. See Cox et al. (2015) for an extended discussion on the comparability between this payoff mechanism and expected utility theory.

[20] Matching sellers and buyers to form individual markets is a popular feature of recent experiments, such as de Haan et al. (2011), Jeitschko and Normann (2012) and Kübler et al. (2008), while a representative example of using a double auction in a pit market is Miller and Plott (1985). The main reason to use matching in the current experiments is to apply random matching among subjects to make each period as close as possible to a one-shot game.

With this method of assignment, we can make sure to get observations for each subject on signaling behavior when selling high- and low-quality products. So, for each parameterization, we can make within-subject comparisons on signaling frequencies (that is, the frequencies of a seller taking the test) given different product qualities. Also, since the population distribution is the key treatment parameter in the Population Treatments, we want to avoid subjects' suspicions on this parameter should a subject experience unusually high or low frequencies of low-quality products.

As soon as the subject matching and product assignment are done, each seller is informed of the quality of her product and is asked to decide whether to test her product. If a seller decides to take the test, a horizontal bar that is divided into success zone and failure zone shows up on the computer screen. A white needle appears and runs across the bar randomly for five seconds. If the needle rests in the failure zone, the product fails the test; if in the success zone, the product passes the test. The failure zone takes up 10% of the bar for the high-type product and 90% for the low type. After all sellers are done with decisions and tests, each product will be put on the market in which it belongs; if a sellers decides to test and the product passes, it will be put on the market with a certificate. The two buyers will be informed whether the product in their market is certified[21] and will then each submit one bid without knowing the other's bid. The product will be sold to the buyer with the higher bid, at a price equal to his own bid. Subjects will then be informed of the two bids and the quality of the product in their market, as well as their own profits in the current period, calculated as in Table 1.4, using Treatment 2 as an example.

After 32 periods of market trading, one period will be randomly chosen to determine all subjects' payoffs.[22] The experiment will then proceed to the second part, in which the subjects need to choose a preferred lottery for each of the 20 lottery pairs we present to them.[23] The 20 lottery pairs are carefully chosen to have good coverage in the Marschak-Machina (MM) triangles (two

[21] Buyers will not be informed of sellers' decisions, and this is commonly known to all subjects.

[22] This is the POR_{PAS} payoff mechanism introduced by Cox et al. (2015), which supports this payoff mechanism as most compatible with Expect Utility theory among all the payoffs tested in the paper.

[23] At the beginning of each session, subjects are informed that there will be a second part in which they can still make money, but not informed of the nature of the tasks in the second part.

Table 1.4: Payoff in Each Period of Treatment 2

Buyer with lower Bid	\$8
Buyer with higher Bid	\$25 − Bid +\$8 if the product has high quality \$10 − Bid +\$8 if the product has low quality
Seller	Product Price − \$5 if she took the test Product Price if she *didn't* take the test

[1] \$8 is paid to all buyers as a non-salient endowment. The purpose is to cover possible losses from overbidding.

pairs will have either common ratios or common consequences), so as to allow robust inference on subjects' risk preferences. Each pair is graphed in its corresponding MM triangle in Figure A.1. The 20 pairs of lotteries are presented to each subject, one pair at a time; the order of the lottery pairs, as well as the positions of the two lotteries in each pair on the computer screen, are individually randomized for each subject. After a subject chooses her preferred lottery in one pair, the chosen lottery will be played out immediately before proceeding to the next choice. After subjects have made all 20 choices, the payoff from one choice will be randomly chosen to determine their payoff in the second part of the experiment.

At the end of each session, subjects are asked to fill out a demographic survey and are paid the sum of their payoffs from both parts of the experiment, plus a show-up fee of \$5. A summary of age, gender, GPA and ethnicity of the subjects in each session can be found in Table B.1.

1.4 Data and Results

In this section, I first describe seller and buyer behavior based on summary statistics and hypothesis tests and then report analysis of treatment effects using different regression specifications. I explore how a subject's risk preference affects her choices of signaling and bidding strategies, and how subjects' strategies as sellers correlate with their own strategies as buyers. I provide only the major tables and graphs, and include the other tables and graphs in the Appendices.

1.4.1 Individual Behavior

For each subject, I create the following four variables to summarize her decisions throughout the 32 periods of market trading: (1) frequency of taking the certification test (referred to as "signaling frequency" hereafter) when selling high-quality products S_H; (2) signaling frequency when selling low-quality products S_L; (3) average bids on certified products B_C; and (4) average bids on non-certified products B_{NC}. For the 48 subjects in each treatment, empirical distributions of signaling frequencies by product quality are shown in Figure 1.1, and empirical distributions of average bids by certification status are shown in Figure 1.4. Also, the means on these variables over the 48 subjects in each treatment are reported in Table 1.5.

Table 1.5: Summaries of Signaling and Bidding Behavior in Each Session

Treatment[1]	Cost	Share of Low Type[4]			Share of Certified Products	Frequency of Signaling[2]			Average Bidding[3] (\$)			Average Price (\$)			Average Price (\$)		
		All μ	Certified μ_1	Non-Certified μ_0		High Type S_H	Low Type S_L	Difference S_{DIFF}	Certified B_C	Non-Certified B_{NC}	Difference B_{DIFF}	Certified	Non-Certified	Difference	Certified	Non-Certified	Difference
1	\$5	75%	4.9%	92.0%	19.3%	83.9%	14.8%	69.1%	16.99	8.36	8.63	19.47	9.04	10.42	24.27	11.20	13.07
2	\$5	50%	1.5%	79.3%	37.1%	79.7%	18.8%	60.9%	16.02	9.06	6.96	18.00	10.35	7.65	24.78	13.11	11.67
3	\$5	25%	0.6%	43.6%	40.9%	59.2%	18.8%	40.5%	17.74	11.86	5.88	19.74	13.92	5.86	24.91	18.47	6.45
4	\$3	50%	2.7%	85.3%	42.2%	88.5%	18.8%	69.8%	19.07	9.08	9.99	21.34	10.09	11.25	24.59	12.21	12.38
2	\$5	50%	1.5%	79.3%	37.1%	79.7%	18.8%	60.9%	16.02	9.06	6.96	18.00	10.35	7.65	24.78	13.11	11.67
5	\$7	50%	2.0%	65.5%	22.8%	49.5%	9.1%	40.4%	13.68	9.32	4.35	15.88	10.51	5.28	24.70	15.17	9.53

[1] In treatments 1, 2 and 4, pooling is unintuitive and separating is the only intuitive equilibrium; in treatments 3 and 5, pooling is intuitive and cannot be refined by Cho-Kreps Intuitive Criterion.

[2] Empirical distributions of S_H, S_L and S_{DIFF} in each treatment can be found in Figure 1.1 and Figure A.2.

[3] Empirical distributions of B_C, B_{NC} and B_{DIFF} in each treatment can be found in Figure 1.4 and Figure A.3

[4] Empirical distributions of μ_0 and μ_1 can can be found in Figure A.4

Starting with seller behavior, I find evidence supporting the separating strategy[24] among sellers, although similar to what has been observed in other signaling experiments, the separation is incomplete. Based on the Wilcoxon paired signed-rank tests on S_H and S_L among the 48 subjects in each treatment, the null hypothesis that sellers take the test at the same frequencies when selling high- and low-quality products can be rejected for all treatments. The alternative hypothesis

[24]This refers to the seller strategy that taking the test when selling high-quality product and not taking the test when selling low-quality ones.

that the sellers take the test more frequently when selling high quality product is supported in all treatments:

$$S_H > S_L \text{ in all treatments, significant at the 1\% level.}$$

However, the empirical distributions of S_H and S_L show different degrees of separation in the five treatments. Figure 1.1 shows the histograms of signaling frequencies by product quality for the 48 subjects in each treatment; the Population Treatments are aligned vertically and the Cost Treatments are aligned horizontally. Treatments with an unintuitive pooling equilibrium are located in the top, left and center panels; in these treatments, there is a clearer separation between the high type's and the low type's distributions: the high-type sellers (the bars with blue solid outlines) are concentrated in the right and closer to the 100% point, while the low-type sellers (the bars with red dashed outlines) are in the left and closer to the 0% point. In contrast, in treatments with an intuitive pooling equilibrium, shown in the right and bottom panels, the distributions of the two types overlap, mainly due to the leftward shift of the high type's distribution, indicating more pooling than in the other three treatments. The means of these distributions are reported in the columns under "Frequency of Signaling" in Table 1.5, and the trends of their changes across treatments are graphed by treatment groups in Figure 1.2.

The observations discussed above are generally supported by two sample Wilcoxon rank-sum tests on S_H and S_L between treatments. Using Treatment 2 (50% and \$5 treatment, the middle panel of Figure 1.1) as the baseline, low-type sellers do not behave differently in other treatments, with one exception: the null hypotheses that ${S_L}^{50\%,\ \$5} = {S_L}^{50\%,\ \$7}$ is rejected at 1% in support of ${S_L}^{50\%,\ \$5} > {S_L}^{50\%,\ \$7}$. In contrast, high-type sellers in treatments with intuitive pooling signal less frequently than high-type sellers in Treatment 2, and high-type sellers in treatments with unintuitive pooling do not show a significant difference:[25]

$${S_H}^{75\%,\ \$5} = {S_H}^{50\%,\ \$5} >^{***} {S_H}^{25\%,\ \$5}$$

[25]The same results hold under the two sample t-tests and Kolmogrov-Smirknov tests with one exception: paired t-tests can reject ${S_H}^{50\%,\ \$3} = {S_H}^{50\%,\ \$5}$ at the 5% level.

$$S_H{}^{50\%,\ \$3} = S_H{}^{50\%,\ \$5} >^{***} S_H{}^{50\%,\ \$7}$$

(*** p< 0.01, ** p< 0.05, * p< 0.1)

For each individual, the difference between the signaling frequencies when selling high- and low-quality products is measured as $S_{DIFF} = S_H - S_L$; the distribution of S_{DIFF} among subjects in each treatment is in Figure A.2, and the means are reported in Table 1.5. Pairwise Wilcoxon rank-sum tests weakly reject the null hypothesis that S_{DIFF} is identically distributed in the three treatments with an unintuitive pooling equilibrium. However, when we compare the treatments with intuitive pooling to Treatment 2, the tests strongly support the alternative hypothesis that the difference in signaling frequencies between high- and low-quality products becomes smaller:[26]

$$S_{DIFF}{}^{75\%,\ \$5} = S_{DIFF}{}^{50\%,\ \$5} >^{***} S_{DIFF}{}^{25\%,\ \$5}$$

$$S_{DIFF}{}^{50\%,\ \$3} = S_{DIFF}{}^{50\%,\ \$5} >^{***} S_{DIFF}{}^{50\%,\ \$7}$$

Conclusion 1: Signaling Behavior *In all treatment, sellers signal more frequently when selling high-quality products. Nevertheless, compared to treatments with unintuitive pooling, the high-quality sellers signal less frequently as the pooling equilibrium becomes intuitive due to an increase in cost or a decrease in the share of low-quality products; as a result, the difference in signaling frequencies between high- and low-quality products is smaller as the pooling equilibrium becomes intuitive.*

As a result of lower signaling frequency among high-type sellers, the percentage of the high (low) type among non-certified products also increases (decreases) in treatments with an intuitive pooling equilibrium. For each individual subject, two variables are created to represent the correlation between the certification status and product qualities that she buys: the share of low-quality products among the certified products she has bought μ_1, and the share of low-quality products among non-certified products μ_0. The means of these two variables over the subjects in each

[26]The same results hold under the two sample t-tests, but Kolmogrov-Smirknov tests can reject $S_{DIFF}{}^{75\%,\ \$5} = S_{DIFF}{}^{50\%,\ \$5}$ at 10% level.

treatment are reported in the columns under "Share of Low Type" in Table 1.5. [27] The distributions of these variables are sharply concentrated around their means, so the histogram is only supplemented as Figure A.4.

The share of low-quality products among certified products μ_1 is very close or equal to 0 in all treatments. The share among non-certified products μ_0 drops *monotonically but not linearly* (solid lines in Figure 1.3) as the population share of low types drops or the cost increases *linearly*. This is caused by a similar non-linear drop of the signaling frequency among high-quality sellers (solid lines in Figure 1.2). The kinks at the mid-point of the lines, where pooling goes from unintuitive to intuitive, suggest the effectiveness of the Cho-Kreps Intuitive Criterion on equilibrium selection. Theoretical predictions on μ_0 under complete pure-strategy separation, that is, when the high type always take the test and the low type never takes the test, are included as the reference lines (dashed lines in Figure 1.3). A bigger distance between the solid line and the dashed line means that the market deviated further from complete separating, and that more high and low types are pooled as non-cerified products.

Conclusion 2: Effectiveness of the Intuitive Criterion *Consistent with the predictions of the Intuitive Criterion, there is dramatically less pooling between the high and low types as non-certified products when the pooling equilibrium is unintuitive. The certified products are predominantly high-quality products in all markets because of the low signaling frequencies and high failure rate among low-quality products.*

[27] The market share of certified products in each treatment is also reported in Table 1.5.

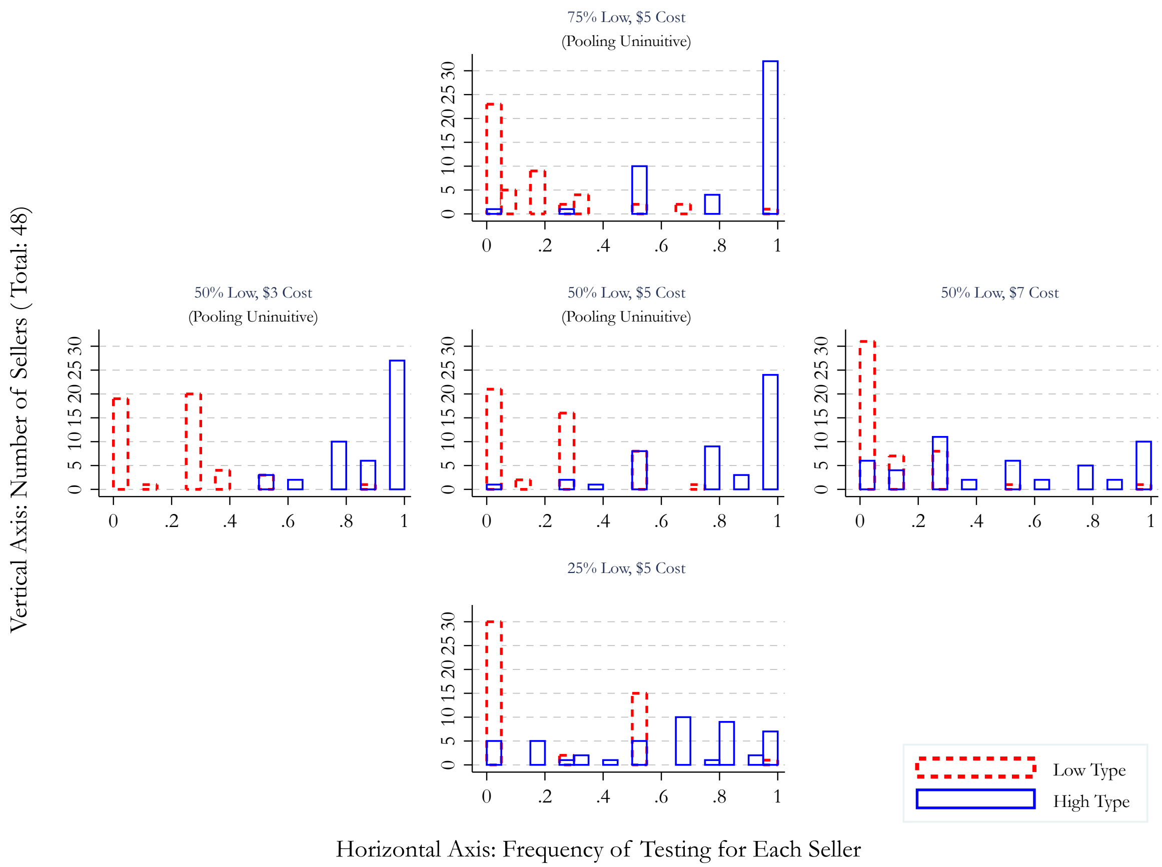


Figure 1.1: Signaling Frequencies per Seller in Each Treatment (by Quality Type)

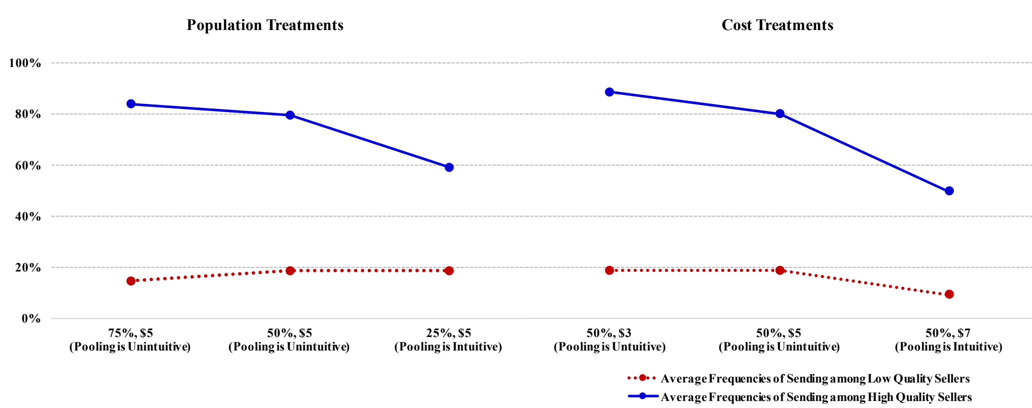
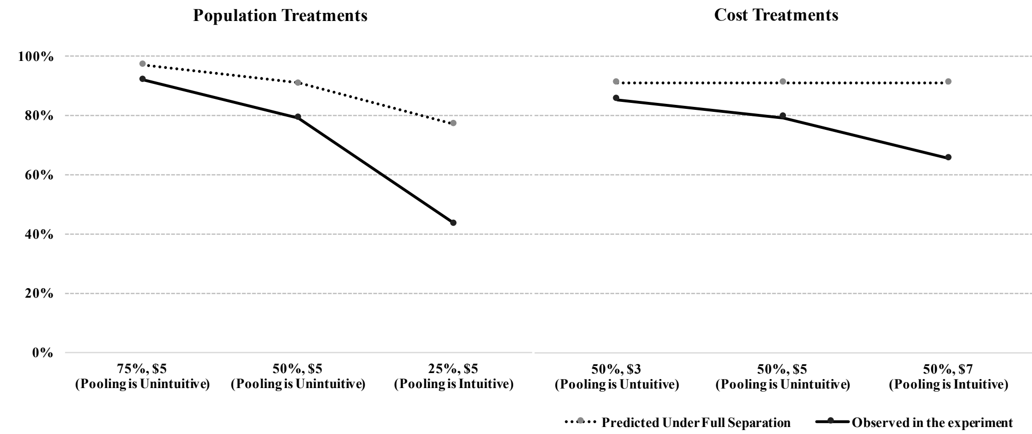


Figure 1.2: Average Frequencies of Signaling in Each Treatment (by Quality Type)



[1] The dashed line is the predicted share of low-quality products among non-certified products, if all sellers with high-quality products take the test and all sellers with low-quality products do not.

[2] The solid line is the observed share of low-quality products among non-certified products in each treatment.

Figure 1.3: Share of Low-Quality Products Among Non-Certified Products

Turning to buyer behavior, Figure 1.4 shows the empirical distributions of average bids on certified and non-certified products over the 48 subjects in each treatment, and the means are reported in the "Average Bidding" columns in Table 1.5. The bidding premium on certified products by each subject – that is, the increase in a subject's average bid on certified, compared to that on non-certified products – is measured as $B_{DIFF} = B_C - B_{NC}$. The distributions are shown in Figure A.3, and the means are also reported in Table 1.5.

The first observation is that bids are generally lower than the expected values of the products given the empirical posterior distributions in each treatment (reported in Table 1.5 and Figure 1.4). This is potentially caused by risk aversion and by the optimal strategy for a subject to bid below her true valuation of the auction object in the first-price sealed-bid auctions[28]. Therefore, rather than comparing bids and prices to the predictions when buyers truthfully reveal their valuations of the object, I focus on comparing bids between certified and non-certified products within each treatment, as well as comparing bids on the same type of products across treatments.

[28] With the same market institution, Kübler et al. (2008) also observe that bids generally do not converge to the predicted value in a pure-strategy separating equilibrium; however, by increasing the number of buyers to three, the bids increase significantly. The conjecture can be tested by switching to alternative institutions, such as a second-price sealed-bid auction or an English auction, in which buyers are predicted to bid their true valuation of the auction object–that is, the certainty equivalents based on their risk attitudes and beliefs (work in progress).

Starting with bidding behavior in Population Treatments, I observe an increasing trend in bids on non-certified products as the population share of the low type decreases: from the top to the bottom panels in Figure 1.4, the distribution of bids on non-certified products shifts to the right, while the distribution of bids on certified products shifts slightly to the left. This pattern is generally confirmed with pairwise Wilcoxon rank-sum tests:

$$B_{NC}{}^{75\%,\ \$5} <^{***} B_{NC}{}^{50\%,\ \$5} <^{***} B_{NC}{}^{25\%,\ \$5}$$

$$B_{C}{}^{75\%,\ \$5} >^{**} B_{C}{}^{50\%,\ \$5} <^{***} B_{C}{}^{25\%,\ \$5}$$

Two factors could drive up the bids on non-certified products as μ decreases: the buyers are updating their posterior beliefs properly in response to more pooling behavior among sellers; and/or they are simply responding to the lower proportion of low types in the prior distribution of the products. In Figure 1.5, I graph the means of B_{NC} across treatments as the solid red line and provide two reference lines: the dashed red line marks the expected value given the posterior distribution of non-certified products observed in each treatment, and the dotted gray line marks the expected value given the posterior in the pure-strategy separating equilibrium. Focusing on the changes in B_{NC} in Population Treatments in the left panel, the slope of the solid red line is somewhere between those of the dashed red line and the dotted gray line, indicating partial but incomplete posterior updating among buyers. That is, buyers do respond to changes in seller's strategies across the treatments, and the increase in B_{NC} is not *solely* due to the change in the share of low types in the prior distribution of the products.[29]

[29] The incomplete posterior updating indicates a mixture of learning modes among buyers, and to fully understand the bidding behavior, a mixture model with different learning models is called for (work in progress).

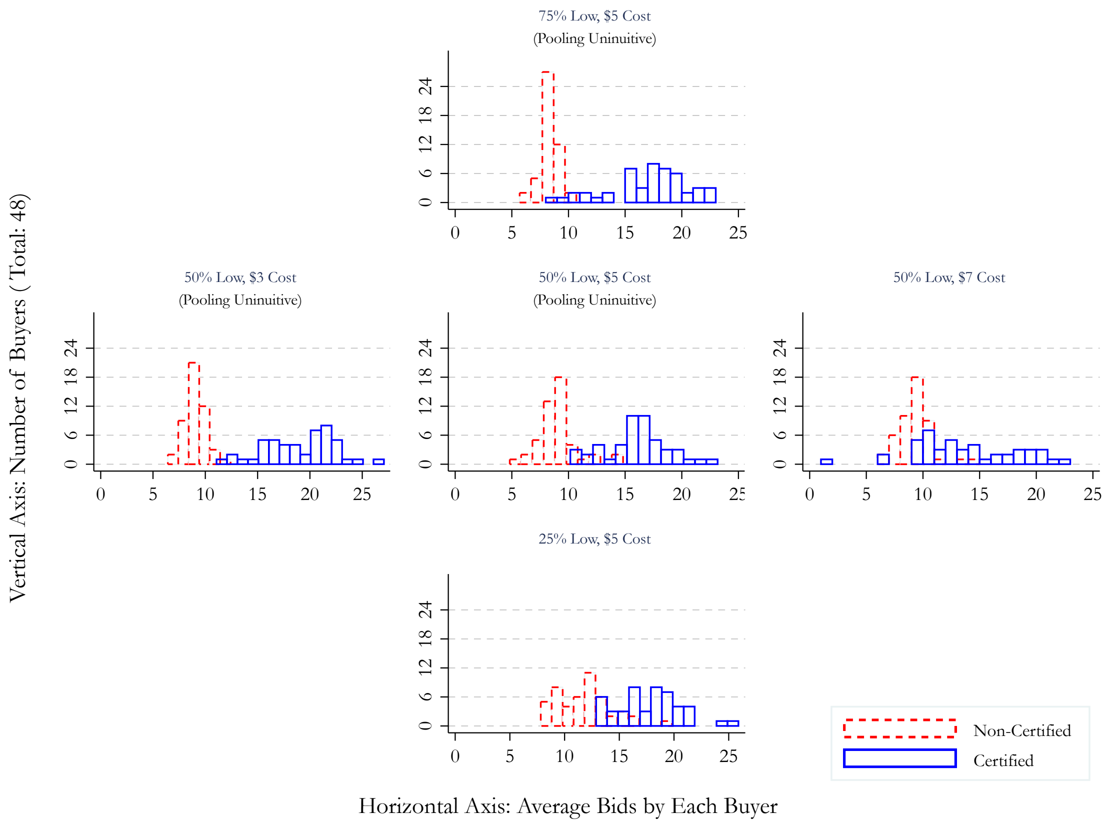
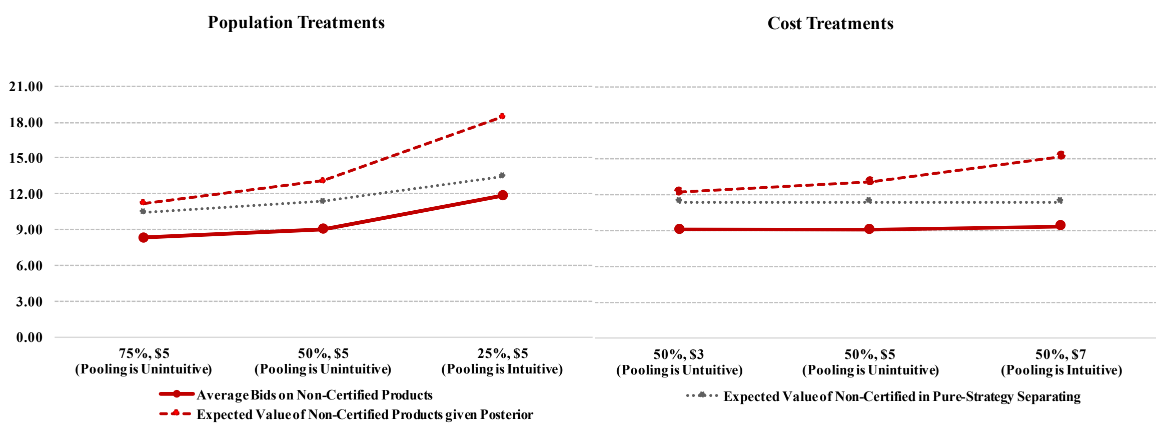


Figure 1.4: Average Bids per Buyer in Each Treatment (by Certification Status)



[1] The dashed red line is the expected value of non-certified products based on the observed posterior in the experiments.

[2] The dotted gray line is the expected value of non-certified products if the market is in pure-strategy separating equilibrium; that is, high types always test and low types never tests.

[3] The solid red line is the mean of bids on non-certified products observed in the experiments.

Figure 1.5: Mean Bids on Non-Certified Products in Each Treatment

Driven mainly by the increase in bids for non-certified products, the bidding difference drops as the share of low types decreases from 75% to 50%, holding the cost constant at $5. This is consistent with the increased pooling behavior among sellers and the decreased separation between high- and low-type products. However, there is no such drop when the share of low types drops further from 50% to 25%. [30]

$$B_{DIFF}{}^{75\%,\ \$5} >^{***} B_{DIFF}{}^{50\%,\ \$5} = B_{DIFF}{}^{25\%,\ \$5}$$

Turning to bidding behavior in Cost Treatments: since the share of low type among non-certified products does not change much when the cost changes from $3 to $5, bids on non-certified products should not change between these two treatments; the share drops when the cost increases further, from $5 to $7, so bids on non-certified products should increase accordingly. These predictions are graphed as the dashed red line in Figure 1.5.[31] The observations are only partly consistent with the above predictions, as is shown in Figure 1.4: consistent with the prediction, as the cost increases (from the left to the right panels), the distribution of B_{NC} does not shift much from the $3 to $5 treatment; however, it also does not shift much from the $5 to $7 treatments, when it should have shifted to the right according to the prediction. I observe similar trends when comparing the means of B_{NC} in these three treatments (the solid red line) in the right panel of Figure 1.5. Pairwise Wilcoxon rank-sum tests cannot reject that the bids on non-certified products are the same in the Cost Treatments:

$$B_{NC}{}^{50\%,\ \$3} = B_{NC}{}^{50\%,\ \$5} = B_{NC}{}^{50\%,\ \$7}$$

Another observation that is inconsistent with posterior updating is that the distribution of bids on certified products shifts to the left when the cost increases from $5 to $7, which should not

[30] While results from Kolmogorov-Smirnov tests are the same, two-sample t-tests can reject $B_{DIFF}{}^{50\%,\ \$5} = B_{DIFF}{}^{25\%,\ \$5}$ at the 10% level.

[31] Since the share of low types in the prior distribution is fixed at 50%, the expected values of non-certified products in *pure-strategy separating equilibrium* is constant and is shown as the dotted gray line.

happen since the posterior distribution of certified products is basically the same between these two treatments. The observations are also backed up by pairwise Wilcoxon rank-sum tests:

$$B_C{}^{50\%,\ \$3} >^{***} B_C{}^{50\%,\ \$5} >^{***} B_C{}^{50\%,\ \$7}$$

Although B_{NC} and B_C do not always change consistently with increased pooling, their joint effect is a dramatic decrease in the bidding premium on certified products B_{DIFF}, which turns out to be consistent with the increase in pooling between high- and low-type products when the cost increases:

$$B_{DIFF}{}^{50\%,\ \$3} >^{***} B_{DIFF}{}^{50\%,\ \$5} >^{***} B_{DIFF}{}^{50\%,\ \$7}$$

Despite the decreased difference in bids when the cost increases and when share of the low type decreases, Wilcoxon paired signed-rank tests reject the null hypothesis that average bids on certified and non-certified products are the same in all five treatments:

$B_C > B_{NC}$ in all treatments, significant at the 1% level.

Conclusion 3: Bidding Behavior *Buyers bid significantly higher on certified products than on non-certified products in all treatments. As the population share of the low type decreases, bids on non-certified products increase, and bids on certified products decrease slightly; as the cost of the certification test increases, bids on non-certified products stay the same, and bids on certified products decrease. As a result, the individual bidding premium on certified products generally drops as μ decreases or as cost increases, consistent with increased pooling between high- and low-quality products (Conclusion 2).*

Prices of certified and non-certified products are also reported in columns "Average Price" of Table 1.5. As one would expect, the prices changes across treatments in a similar pattern to bids: the price premium of certified products (defined as the price difference between certified and non-certified products) drops as the percentage of the low type decreases or as the cost of the certification test increases. With regard to seller behavior, Table 1.6 shows, for each type, the expected payoff of sellers when they take the test and when they don't, given the average prices in each treatment. The signaling frequency of high-type sellers is, indeed, positively correlated with their expected gain from taking the test (the "Difference" column).[32]

An interesting question to ask, given the role switching design of this experiment, is whether an individual's behavior as a seller correlates with her own strategy as a buyer. A simple way to answer this question is to find the correlation coefficients between the differences in signaling frequencies S_{DIFF} and bidding premiums B_{DIFF} among the 48 subjects in each treatment. The results are reported in Table 1.7. If the beliefs about others' strategies come from a subject's own perception of the game and choice of strategies, we will expect B_{DIFF} to be positively correlated with S_{DIFF} within each treatment. Such correlations are found in all treatments except for Treatments 1. I further decompose these correlations in terms of the bidding behavior on certified and non-certified products, and find that they mostly caused by positive correlation-ships between B_C and S_{DIFF}.

Table 1.6: Correlation between Price Premium and Seller Decision in Each Treatment

Treatment	Cost	Share of Low Type	Average Price ($)			Expected Payoff of High Type ($)			Signaling FrequencyHigh Type S_H	Expected Payoff of Low Type ($)			Signaling Frequency Low Type S_L
			Certified	Non-Certified	Difference	Test	No	Difference		Test	No	Difference	
1	$5	75%	19.47	9.04	10.42	13.43	9.04	4.38	83.9%	5.08	9.04	-3.96	14.8%
2	$5	50%	18.00	10.35	7.65	12.24	10.35	1.89	79.7%	6.12	10.35	-4.23	18.8%
3	$5	25%	19.74	13.92	5.86	14.16	13.92	0.24	59.2%	9.51	13.92	-4.42	18.8%
4	$3	50%	21.34	10.09	11.25	17.21	10.09	7.12	88.5%	8.22	10.09	-1.88	18.8%
2	$5	50%	18.00	10.35	7.65	12.24	10.35	1.89	79.7%	6.12	10.35	-4.23	18.8%
5	$7	50%	15.88	10.51	5.28	8.34	10.51	-2.17	49.5%	4.04	10.51	-6.46	9.1%

Table 1.7: Self-Correlation Between Subjects' Signaling and Bidding Strategies

Low Type	Cost	$B_{DIFF} \times S_{DIFF}$	$B_{NC} \times S_{DIFF}$	$B_C \times S_{DIFF}$
75%	$5	0.093	-0.058	0.080
50%	$5	0.269 *	-0.042	0.245 *
25%	$5	0.485 *	-0.408 *	0.313 *
50%	***$3***	0.346 *	-0.089	0.334 *
50%	***$7***	0.571 *	0.130	0.564 *

* Significant at 10% level

[32]However, the table also shows that sellers are taking the test with positive frequencies even when the expected payoffs are negative. This indicates that they are either risk loving when they make signaling decisions, or that sellers do not fully adjust to the market prices in their investment decisions.

To see how subjects learn to play the game, as well as how risk preferences affect their decisions as buyers and sellers, I report random effects panel regression results on signaling and bidding decisions in Table 1.8, with errors clustered at the subject level. Using observations within each treatment, I run logit regressions on sellers' decisions of whether or not to take the certificate test in each period with three sets of explanatory variables, and I arrange the results by treatment groups for ease of comparison. The models with product type as the only explanatory variable (first three columns) are consistent with the hypothesis tests and Conclusion 1: the coefficients are all positive at the 1% significance level, and generally decrease as the percentage of low types drops in the Population Treatments or as the cost increases in the Cost Treatments. To directly test if the changes in the coefficients are significant, I run panel regressions on subjects signaling behavior separately for high and low quality products, with 50%, \$5 session as baseline session and dummies for other treatments, and with errors clustered at the subject level. The results are reported as Table 1.10 and further confirm the discussions above.

When I add the number of current periods interacted with product type to the set of control variables (the middle three columns), I find that, subjects learn to signal more often when selling high-type products in the three treatments with an unintuitive pooling equilibrium, but do not in treatments with an intuitive pooling equilibrium.[33] The last three columns in Table 1.8 report the random effects panel regression results, which include the number of safer lotteries[34] that subjects choose in the 20 lottery pairs as the control variable for risk preference. The signs of the coefficients are significantly negative only in two treatments, as predicted in the discussion on

[33]Given the experimental design as demonstrated in Table 1.3, I compare the signaling behavior of subjects who played as seller in the first and last blocks. I calculate each subject's signaling rates S_H, S_L and S_{DIFF} in the first and last block, and report the means over all subjects within each treatment in Table B.2. To test for the differences in the distributions of these variables between the two blocks, I perform Wilcoxon Signed-Rank tests on these variables. I find that when selling high type products (S_H), subjects signal more often in the last block than in the first block in two of the three treatments with unintuitive pooling, but not so in either of the two treatments with intuitive pooling. I find no significant difference in signaling behavior between the first and last blocks when selling low type products (S_L), except for in Treatment 5. I also find that, compared to the first block, the difference in signaling rates $S_H - S_L$ is greater in the last block in all three treatments with unintuitive pooling, but not so in the two treatments with intuitive pooling.

[34]The safer lottery in a pair refers to the one positioned left in the MM triangle. Under the Expected Utility Theory, indifference curves are linear in the MM triangles. Also, the more risk averse an individual is, the steeper her indifference curve, and therefore, the more likely she chooses the lottery positioned left from a given lottery pair.

risk attitudes in the Model Section: the more risk-averse a subject is, the less likely it is that she will choose to take the test. However, risk attitudes do not significantly affect sellers' decisions in other treatments.

Similarly, I fit three specifications of linear regression models on bids in each treatment, which are reported in Table 1.9. The first three columns are models that include certification status as the only explanatory variable, and the results are consistent with the hypothesis tests and Conclusion 3.[35] Then, I add the interaction term between period and the product type (the middle three columns), and the results show an increasing trend in the bids. Also, subjects learn to bid more on certified products in Treatments 2 (50%, \$5) and 4 (50%, \$3), while learning to bid more on non-certified products in all other treatments. I do not observe any coefficients to be significantly negative. This suggests generally increased competition among buyers.[36] Except in Treatment 3 (25%, \$5), there is no significant changes in bidding behavior due to risk attitudes after adding number of safe choices in the linear specification.[37]

As robustness check, I also run regressions on the signaling and bidding behavior, pooling observations from all five treatments (Treatment 2 as the baseline, errors clustered at the subject level) and using four sets of explanatory variables. Tables B.4 and B.5 report the full regression results with controls on demographic variables. Most demographic variables do not have any significant effects in subjects' behavior, with only a handful of exceptions.

[35] To directly test if these changes, I also report panel regressions on subjects bidding behavior separately for certified and non-certified products in Table 1.11, with 50%, \$5 session as baseline and dummies for other treatments, and with errors clustered at the subject level.

[36] I calculate average bids of each subject in the first and last blocks of the experiment on certified products B_C, on non-certified products B_{NC} and also on the bidding premium for certified products B_{DIFF}, and Table B.3 reports the means of these three variables in the first and last blocks of each treatment, as well as the differences between the two blocks. Within each treatment, I test for the differences on the distributions of these three variables between the first and last blocks using Wilcoxon Rank-Sum tests. Compared to their counterparts in the first block of experiments: average bids on certified products B_C are significantly higher in the last blocks of all treatments; average bids on non-certified products B_{NC} are also significantly higher in the last blocks of all treatments except for in Treatment 4; and the bidding premium $B_C - B_{NC}$ are higher in the last blocks of the treatments where pooling is unintuitive, all significant except for Treatment 1.

[37] A possible cause of this result is that when subjects make decisions in the lottery task, they already know their payoff from the first part of the experiment, so the changes in wealth level might lead to changes in risk attitude. The insignificance could also possibly be caused by the linear specification between bids and number of bids, and a structural model based on the optimal bidding function might be needed to look further into this matter.

Table 1.8: Panel Regression on Individual Signaling Behavior by Treatment

1=Signal	Population Treatments								
	75% Low	50% Low	25% Low	75% Low	50% Low	25% Low	75% Low	50% Low	25% Low
High	4.283***	3.296***	2.579***	3.078***	2.337***	2.018***	3.008***	2.348***	2.021***
	(0.383)	(0.276)	(0.305)	(0.608)	(0.479)	(0.556)	(0.603)	(0.479)	(0.556)
Period				-0.0245	-0.0248	-0.0353	-0.0240	-0.0245	-0.0353
				(0.0194)	(0.0200)	(0.0279)	(0.0190)	(0.0199)	(0.0279)
High x Period				0.0822**	0.0624**	0.0358	0.0835**	0.0621**	0.0358
				(0.0354)	(0.0268)	(0.0307)	(0.0353)	(0.0268)	(0.0307)
Number of Safe Choices							-0.144***	-0.0792*	-0.0389
							(0.0538)	(0.0450)	(0.0610)
Constant	-2.223***	-1.745***	-2.060***	-1.850***	-1.368***	-1.507***	-0.162	-0.527	-1.074
	(0.263)	(0.219)	(0.335)	(0.393)	(0.378)	(0.534)	(0.713)	(0.595)	(0.859)
1=Signal	**Cost Treatments**								
	$3	$5	$7	$3	$5	$7	$3	$5	$7
High	3.463***	3.296***	2.849***	2.438***	2.337***	2.497***	2.430***	2.348***	2.496***
	(0.274)	(0.276)	(0.311)	(0.483)	(0.479)	(0.551)	(0.483)	(0.479)	(0.551)
Period				-0.0107	-0.0248	-0.0528*	-0.0109	-0.0245	-0.0528*
				(0.0172)	(0.0200)	(0.0279)	(0.0171)	(0.0199)	(0.0279)
High x Period				0.0686**	0.0624**	0.0269	0.0690**	0.0621**	0.0270
				(0.0282)	(0.0268)	(0.0318)	(0.0282)	(0.0268)	(0.0318)
Number of Safe Choices							0.0378	-0.0792*	0.0133
							(0.0397)	(0.0450)	(0.0549)
Constant	-1.438***	-1.745***	-2.876***	-1.270***	-1.368***	-2.093***	-1.713***	-0.527	-2.237***
	(0.181)	(0.219)	(0.333)	(0.330)	(0.378)	(0.500)	(0.576)	(0.595)	(0.780)

For each regression: Number of Observations 512, Number of Subjects 48

Robust Standard errors in parentheses, errors clustered at the subject level.

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

Table 1.9: Panel Regression on Individual Bidding Behavior by Treatment

Bids	**Population Treatments**								
	75% Low	50% Low	25% Low	75% Low	50% Low	25% Low	75% Low	50% Low	25% Low
Certified	8.962***	6.979***	5.973***	8.195***	4.958***	5.387***	8.195***	4.960***	5.389***
	(0.477)	(0.431)	(0.607)	(0.941)	(0.923)	(0.845)	(0.941)	(0.924)	(0.845)
Period				0.0483***	0.0304	0.0773***	0.0483***	0.0304	0.0777***
				(0.0125)	(0.0238)	(0.0213)	(0.0125)	(0.0239)	(0.0213)
Certified x Period				0.0418	0.116***	0.0320	0.0418	0.116***	0.0310
				(0.0343)	(0.0436)	(0.0332)	(0.0343)	(0.0436)	(0.0333)
Number of Safe Choices							0.0128	0.0175	0.166**
							(0.0374)	(0.0608)	(0.0733)
Constant	8.355***	9.064***	12.05***	7.566***	8.579***	10.79***	7.416***	8.389***	8.945***
	(0.118)	(0.252)	(0.384)	(0.286)	(0.520)	(0.528)	(0.465)	(0.953)	(1.027)
Bids	**Cost Treatments**								
	\$3	\$5	\$7	\$3	\$5	\$7	\$3	\$5	\$7
Certified	10.05***	6.979***	5.232***	8.184***	4.958***	5.047***	8.180***	4.960***	5.046***
	(0.535)	(0.431)	(0.613)	(0.865)	(0.923)	(0.776)	(0.867)	(0.924)	(0.776)
Period				-0.00455	0.0304	0.0748***	-0.00460	0.0304	0.0748***
				(0.0115)	(0.0238)	(0.0153)	(0.0115)	(0.0239)	(0.0153)
Certified x Period				0.110***	0.116***	0.0204	0.111***	0.116***	0.0203
				(0.0370)	(0.0436)	(0.0302)	(0.0370)	(0.0436)	(0.0302)
Number of Safe Choices							0.0273	0.0175	-0.0378
							(0.0582)	(0.0608)	(0.0602)
Constant	9.130***	9.064***	9.361***	9.208***	8.579***	8.098***	8.889***	8.389***	8.512***
	(0.158)	(0.252)	(0.220)	(0.291)	(0.520)	(0.363)	(0.789)	(0.953)	(0.798)

For each regression: Number of Observations 1024, Number of Subjects 48

Robust Standard errors in parentheses, errors clustered at the subject level.

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

Table 1.10: Regression Analysis on Signaling Behavior

Signal	Low Quality Products			High Quality Products		
	(1)	(2)	(3)	(5)	(6)	(7)
Treament 1 (75%, $5)	-0.254	-0.255	-0.188	0.285	0.286	0.281
	(0.315)	(0.318)	(0.318)	(0.361)	(0.361)	(0.362)
Treament 3 (25%, $5)	0.0277	0.0278	0.0775	-0.887***	-0.889***	-0.892***
	(0.326)	(0.327)	(0.331)	(0.306)	(0.306)	(0.308)
Treament 4 (50%, $3)	0.233	0.234	0.346	0.603*	0.604*	0.597*
	(0.290)	(0.291)	(0.303)	(0.321)	(0.322)	(0.325)
Treament 5 (50%, $7)	-0.715**	-0.718**	-0.725**	-1.374***	-1.377***	-1.378***
	(0.352)	(0.352)	(0.339)	(0.316)	(0.317)	(0.317)
Period		-0.0206***	-0.0211***		0.0103	0.0103
		(0.00733)	(0.00717)		(0.00745)	(0.00745)
Number of Safe Choices			-0.0850**			0.00558
			(0.0344)			(0.0278)
Constant	-1.600***	-1.272***	-0.393	1.343***	1.175***	1.116***
	(0.209)	(0.235)	(0.410)	(0.239)	(0.272)	(0.403)

For each regression: Number of Observations: 1280

Robust standard errors in parentheses

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

Table 1.11: Regression Analysis on Bidding Behavior

Bids	Non-Certified Products			Certified Products		
	(1)	(2)	(3)	(4)	(5)	(6)
Treament 1 (75%, $5)	-0.799***	-0.806***	-0.813***	1.298**	1.278**	1.212*
	(0.277)	(0.276)	(0.272)	(0.641)	(0.638)	(0.645)
Treament 3 (25%, $5)	2.912***	2.908***	2.908***	1.886***	1.910***	1.889***
	(0.482)	(0.480)	(0.479)	(0.642)	(0.640)	(0.633)
Treament 4 (50%, $3)	-0.0157	-0.0184	-0.0240	3.158***	3.180***	3.131***
	(0.289)	(0.290)	(0.287)	(0.651)	(0.655)	(0.659)
Treament 5 (50%, $7)	0.113	0.0800	0.0788	-1.121	-0.920	-0.898
	(0.334)	(0.331)	(0.331)	(0.775)	(0.769)	(0.773)
Period		0.0479***	0.0479***		0.112***	0.112***
		(0.00796)	(0.00796)		(0.0157)	(0.0157)
Number of Safe Choices			0.00764			0.0938
			(0.0282)			(0.0573)
Constant	9.146***	8.372***	8.288***	16.07***	14.15***	13.11***
	(0.250)	(0.299)	(0.462)	(0.431)	(0.551)	(0.818)
Observations	3,456	3,456	3,456	1,664	1,664	1,664
R-squared	0.172	0.194	0.194	0.090	0.136	0.141

Table 1.10 reports the regressions on signaling frequencies for low- and high-type sellers, respectively. The treatment effects are generally consistent with those found from the hypotheses tests and panel regressions on signaling behavior, whether or not we add period or the number of safe choices as control variables. In contrast to the panel regression results of each treatment, I find here that subjects learn to signal less often when they sell low-quality products, and I find no learning effects when they sell high-quality products. The coefficients of the risk attitudes variable suggest that more risk averse sellers signal less frequently when selling low type products, but not so when selling high type products.

Table 1.11 shows the regression results for bids on non-certified and certified products, respectively. The treatment effects are also consistent with the results from hypotheses tests and panel regressions. The coefficients of the period are significantly positive in all treatments and are greater for bids on certified products than on non-certified products. This is consistent with the finding from panel regressions that competition among buyers tends to increase as the experiment continues, and more so among buyers of certified products.

1.5 Concluding Remarks

In pursuing an education degree, people face different risks of dropout due to different abilities. Inspired by this observation, I developed a general signaling model by introducing type-dependent noises into a simplified Spence model. Previous research on stochastic signaling models commonly assumes homogeneous noises and relies on type-dependent costs to generate a separating equilibrium. In contrast, the current paper highlights the role of type-dependent noises in inducing separation, when signaling costs cannot be effectively differentiated between different types of senders.

Both the theoretical analysis and the experimental results have shown that, if the more desirable type of senders have better control over the signal-generating process, a separating equilibrium might exist, even if sending the signal is equally costly regardless of sender type. The paper also provide lab evidence that supports the effectiveness of the Cho-Kreps Intuitive criterion – in treatments where pooling is unintuitive, I observe dramatically less pooling in both senders' and receivers' behavior. The paper also contributes to the literature on signaling experiments with the discussions and tests on how subjects' risk preferences affect their strategies in stochastic signaling games.

Developed in the context of job market signaling, the paper aims to bring new insights to settle the contest between human capital and education signaling models in explaining wage differentials associated with higher education. If the risk of dropout weighs in on people's enrollment decisions and can cause a similar self-selection process discussed by Spence, then we should pay due attention and gather data on this variable when we decompose the role of education in signaling and improving productivity. For instance, a testable hypothesis from the model is that, between two programs that offer similar curricula and training, we would expect graduates from the one with more stringent degree requirements to have a bigger premium on wages due to more significant signaling effects.

Admittedly, the simplicity of the model developed in the current paper allows for a cleaner test but also limits its applications in addressing more complex questions. To make the model more realistic, an interesting extension for future research would be to develop a dynamic model by breaking down the graduation process and examining the costs and dropout risks that students face each year. In a different vein, more experiments can be run to determine the effects of market institutions on equilibrium, specifically, on the level of product prices and the price premium of certified products.

Chapter 2

A Bayesian Mixture Model for Decisions under Risk[1]

2.1 Introduction and Literature Review

How do people make decisions under risk? This is an important question given the prevalence of risks in economic and social situations, and numerous studies are devoted to answer this question from different perspectives – how to collect the most informative data, what are the alternative models, and which models best characterize the behavior pattern revealed by the data. In the quest for the answer, the importance of allowing for heterogeneity in modeling individuals' risk preferences and decision processes has become apparent. We apply the Bayesian Econometric methods to the structural estimation of individual risk preferences, and estimate a mixture of different decision models to allow for heterogeneity in preference types and response modes within a group of decision makers.

We focus on the analysis of experimental data, which is one of the most important data sources to shed light on this question. With carefully designed procedures, researchers recruit individuals into the laboratory, collect decision samples from them on decision problems that involve known risks, and then use statistical methods to infer the underlying decision processes that have generated the observed decisions. The risky situations that subjects face in the laboratory are induced by researchers, and therefore we have complete information on the parameters of the decision environment (probabilities, outcomes, etc); in contrast to experimental data, observational data

[1] Coauthored with Dr. Glenn W. Harrison and Dr. Rusty Tchernis. I thank Dr. Glenn W. Harrison and Dr. Todd Swarthout for providing the observed data used in Section 2.5.

typically involves naturally occurring risks, leading to uncontrolled potential confounds. Apart from the benefit of allowing the *researchers* to know about the relevant decision parameters, the researchers can control what the decision makers know about the decision environment, and the researchers thus have knowledge on what the *decision makers* know in making these decisions. Finally, the decision tasks used in experiments targeted at inferring risk preferences are typically simple and non-interactive, which eliminates potential confounding of strategic or altruistic considerations in decisions.

We focus on the analysis of two popular theories, Expected Utility Theory (EUT) and Cumulative Prospect Theory (CPT), among the many contending theories on preferences under risks.[2] EUT has been the standard decision theory under risk since formalized by Neumann and Morgenstern (1944) as a core component of game theory, and Schoemaker (1982) has provided a comprehensive review on the variants and early empirical evidence. One of the key assumptions underlying EUT, the Mixture Independence Axiom, has been under attack on the basis of choice anomalies reported by experimenters. In response, theorists have developed alternative preference theories to EUT in an effort to accommodate these stylized anomalies, as reviewed in Starmer (2000). Among the many alternative theories, one popular contender is CPT developed in Tversky and Kahneman (1992), which is based on the original Prospect Theory proposed by Kahneman and Tversky (1979). In contrast to EUT, CPT can accommodate phenomena of framing effects, loss aversion, likelihood insensitivity and pessimism/optimism, and has been widely applied.

Our focus is the statistical inference methods linking data and theories. In one direction, many theories are inspired by decision patterns revealed in experiments, and in the other direction, the validation of a theory also calls for a careful examination of how well it fits the experimental data. It seems intuitive that different people think and behave differently and therefore can be characterized by different decision models. Moreover, both EUT and CPT typically assume that individuals can have different risk preferences. It's natural to expect some subjects to be better

[2]Other popular alternative theories include Rank Dependent Theory Quiggin (1982), Dual Theory Yaari (1987) and Disappointment Aversion Theory Gul (1991).

characterized by EUT, and others by CPT.[3] Rather than forcing one decision model to explain all subjects' behavior, we need a mixture model to allow for the co-existence of different decision models generating the population data.

Allowing the observations to be generated by different decision models, three studies have used mixture modeling methods in the structural estimation of risk preferences.[4] Harrison and Rutström (2009) estimate a mixture between EUT and the original Prospect Theory, and address individual heterogeneity within each theory by allowing preference parameters to be a linear function of demographics. Conte et al. (2011) estimate a mixture between EUT and Rank Dependent Utility (RDU), and adopt a random parameter approach to address heterogeneity within each decision type, as well as the mixing probability. Bruhin et al. (2010) use finite mixture methods to classify individuals into different decision modes (either EUT or RDU); similar to Wilcox (2008), they look into individual heterogeneity through subject and task specific noise specifications; in contrast to Harrison and Rutström (2009) and Conte et al. (2011), they do not allow parameters to differ across individuals. Another important distinction of Harrison and Rutström (2009) is that they assume each individual use both types of decision modes and switch between EUT and CPT across decisions, while the other two studies classify each subject as only using one type of decision mode in all decisions.

Extending these studies, we introduce Bayesian Econometric methods into the estimation. As the first step of the research project, we develop and test the Bayesian algorithms with pooled data (simulated or observed) in the current chapter. We specify parametric functionals for EUT and CPT, and estimate the parameters based on these specifications. To simplify, we assume there is no heterogeneity in the functionals or parameters within the same preference type.[5] We base our analysis on the data and experiments reported in Harrison and Swarthout (2016), and use

[3]It's also undeniable that subjects can also be characterized by other theories. However, we choose EUT and CPT as a starting point for the purpose of illustrating the statistical methods.

[4]Mixture modeling has also been applied to the analysis of other types of experimental data since the 1990s: see El-Gamal and Grether (1995), Stahl and Wilson (1995), Haruvy et al. (2001), Hurley and Shogren (2005), and Bardsley and Moffatt (2007).

[5]We explore the heterogeneity within the same type in a separate study, in which we take a Hierarchical Bayesian approach.

both simulated data and observed data. We use simulated data with known underlying preference functionals and parameters in order to test if our algorithms can recover the true parameters, and then we apply the algorithms to analyze the actual choice data from Harrison and Swarthout (2016).

We are not the first to apply Bayesian Econometric methods to the analysis of experimental data. Houser et al. (2004) uses Bayesian algorithms (Gibbs sampler) to classify subjects' behavior in a complicated dynamic decision problem into one of three types of decision rules. Although the decision tasks are different, their inference task is similar to ours since the type of each subject is unknown, and the parameters that characterize each type are also unknown. As is pointed out by McLachlan and Peel (2000) which provided a modern survey of developments in mixture models, Maximum Likelihood Estimation (MLE) method can be problematic in this type of inference, since the likelihood tends to be ill-conditioned and have multiple local maxima, which makes the task of finding the global maximum relatively difficult. In contrast, Bayesian methods draw inferences free of the need to maximize the likelihood, and are "less sensitive to problems created by ill behaved likelihood surfaces (e.g., local maxima) than is ML." (Houser et al. (2004; p.793)). Nilsson et al. (2011) is the most relevant to our study in terms of the experimental decision tasks: they focus on the estimation of individual risk preferences and take a Hierarchical Bayesian modeling approach. However, they restrict their analysis to fitting the data to CPT rather than a mixture model of different response modes[6].

Using the algorithms developed in the current chapter, we can continue the research project in two directions.

In the first direction, we estimate the mixture model for each individual based on her choice data, through which we can both acquire estimates on the preference parameter and naturally determine her preference type by examining the mixture probability.[7] The joint estimation of many

[6] Murphy and ten Brincke (2017) also estimated a hierarchical model of CPT, but with a different estimation method, the "Hierarchical Maximum Likelihood Estimation" method, which uses a random coefficient approach estimated with the Maximum Simulated Likelihood method.

[7] A different approach to individual level structural estimation, illustrated by Hey and Orme (1994) and Harrison and Ng (2016), is to estimate non-mixture models for each subject, and then use either likelihood ratio tests or Akaike Information Criterion to determine which model better characterizes her decision type. This approach would be

parameters for each individual is enhanced by the use of Bayesian econometric methods, since the sample size we can collect from each individual is typically small,[8] which can potentially lead to ill-conditioned likelihoods. With ill-conditioned likelihoods in the high dimensional parameter spaces, the search for the global maximum can be very difficult, or generate estimates that are simply unrealistic. With the Bayesian framework, we can solve the latter problem by including non-diffuse, informative priors: in estimating the parameters for each individual, we can use the population distributions of these parameters in the same subject pool as the prior distributions. Then we can apply algorithms to get representative samples from the posterior distributions of parameters, an approach that could potentially be more feasible than the global maximum searching algorithms that rely on derivatives, such as the Newton-Raphson (NR) algorithm. More appropriately, if we believe parameters might be differently distributed among different subgroups of the same subject pool, we may use the sub-population level parameter distributions conditioned on the individual's observable characteristics as the prior for her estimation.

With appropriate priors as discussed above, if a subject's decisions are very noisy the posteriors will mainly reflect the knowledge from the priors; if not, the posteriors will reflect a combination of the prior knowledge from population behavior and the information from individual choices. Compared to MLE, we expect this approach to generate more robust and realistic estimates, especially when a subject's decisions are inherently quite noisy. Another advantage of Bayesian econometric method is that, while hypothesis tests on MLE estimates rely on the asymptotic properties of the estimators suitable for large samples, hypothesis tests on Bayesian estimates do not rely on such properties. For instance, if we are to test linearity of a subject's utility function at the 90% confidence level, we will just need to find if the linear parameter value lies between the 5% and 95% quantiles of the posterior sample.

---

problematic when a subject inherently uses two or more decision modes when making the observed decisions, as discussed in Andersen et al. (2014b).

[8] One can increase the number of decision tasks, but can presumably only do so to a limit without increasing decision noise. Experiments in which subjects make between 60 and 100 lottery choices, to facilitate individual-level estimation, are now common: see Harrison and Ng (2016) and Harrison and Swarthout (2016).

In the second direction, we use the Bayesian econometric framework to analyze two-stage experimental data, as illustrated by the design of experiments reported in the first chapter. To recap, the experiment has two stages: in the main stage, subjects play the signaling games, and in the preference elicitation stage, subjects complete binary lottery tasks. To extend the analysis, we hypothesize the underlying decision process with a structural decision model, and estimate its parameters for each subject based on her observed signaling decisions. In doing so, we will need to jointly estimate a subject's risk preference and a noise parameter that measures the predictive power of the structural decision model. The number of signaling decisions are typically 8 or 16 for each subject, which makes the joint estimation difficult, and sometimes infeasible, for the same reasons as discussed above.

One solution is to simply substitute the point estimates of a subject's risk preference parameter from her binary lottery choices in the structural signaling model, and estimate only the noise parameter. We propose a better solution: estimating the full distribution of risk preference parameters from binary lottery tasks, and use it as the prior in the estimation of the structural model on signaling decisions. We can then update this prior with the subject's signaling decisions, which allows for the possibility that her risk preferences are related, but not identical, in the two stages. In contrast, the first solution fixes the risk preference parameter at the point estimate, and the validity of the estimation for the noise parameter implicitly rests on the assumption that her risk preferences are identical in these two stages. As the two-stage design is becoming a popular feature among recent experiments, the proposed solution can be used to address data analysis in these experiments similarly. For instance, Andersen et al. (2008) design experiments to jointly estimate time and risk preference, in which subjects respond to four risk aversion tasks and six discount rate tasks.[9] In order to infer a subject's discount rate from her decisions in the discount rate tasks, one needs to know her risk attitude. An application of the proposed solution would be to estimate a subject's risk preference parameter from the risk aversion tasks, and use it as the prior in the joint estimation of risk preference and time discounting rate from the discount rate tasks.

[9] They undertake full joint estimation with pooled data, but not at the individual level. With a different design, and more choices per individual, they undertake joint estimation at the individual level in Andersen et al. (2014a).

In the next section we describe the decision tasks that are used to collect or simulate data for our analysis. In Section 2.3 we outline our specifications on the preference and decision models, describe the simulated data, and then derive the posterior distribution of the parameters. In Section 2.4 we introduce the Markov Chain Monte Carlo (MCMC) algorithms used to sample from the posterior and present the estimation results. In Section 2.5 we use the algorithms to estimate the mixture model using observed decision data from real subjects; and in Section 2.6 we draw some conclusions.

2.2 Data

There are many types of decision samples that experimenters can collect from subjects to infer their risk attitudes, as reviewed in Harrison and Rutström (2008) and Charness et al. (2013). We focus on Binary Lottery Tasks, where the experimenter presents each subject with series of lottery pairs, and each subject needs to choose from each pair of lotteries which lottery he or she prefers. We choose Binary Lottery Tasks because, when the lottery pairs are properly chosen, subjects' responses to these tasks can be very informative for the purpose of the structural estimation of their risk preferences. We simulate subjects' choices using the 100 lottery pairs in the experiments reported in Harrison and Swarthout (2016),[10] which were themselves inspired by Loomes and Sugden (1998) and Conlisk (1989). Figure 2.1 presents these choice pairs in the Marschak-Machina (MM) triangle. Each lottery involves a maximum of three unique prizes that have positive probabilities and can be represented by a point in the probability space with three axes corresponding to the probabilities of prizes. The MM triangle is the projection of the probability space onto the plane spanned by the axes representing the probabilities of the best and worst prizes. Each MM triangle provides a context of three prizes, and two lotteries in each pair are connected with a straight line in the corresponding MM triangle. As an important design feature,

[10] Harrison and Swarthout (2016) used a different set of lottery pairs for undergraduate students and MBA students, by scaling up prizes for the latter. We are using the lottery pairs for undergraduate students.

lottery pairs are grouped in terms of the slopes of their connecting lines within their corresponding MM triangles in Figure 2.1: under EUT, indifference curves within the MM triangle are linear.

We summarize the many desirable features in the design of these lottery pairs here, to make the point that subjects's choices (simulated or observed) in these tasks are informative about the underlying risk preferences and are suitable for the purpose of structural estimation to be performed. For an expanded discussion and numerical examples, see Harrison and Swarthout (2016; §1). In order to test for the validity of EUT, two pairs of lottery pairs have common consequences and allow us to observe the existence of Allais Paradox or the lack thereof (the first two panels in Figure 2.1). The other 96 lottery pairs can be categorized into 12 groups (the other 12 panels in Figure 2.1), and lotteries in each group have the same slope of the connecting lines within the same MM triangle, which allows us to test for Common Ratio effects. Moreover, there is a lot of variation in the slopes of the connecting lines across different groups of lottery pairs, which allows for powerful tests of EUT, irrespective of the risk attitudes of subjects under EUT (Loomes and Sugden (1998)). Harrison et al. (2007) refers to this design as a "complementary slack experimental design:" the risk attitude of a given EU subject is characterized by the slope of linear indifference curves in an MM triangle, so if we observe common ratio effects in one or two groups of lotteries, and if we attribute them to noises and conclude that the indifference curves of the subject might have a slope the same as (or close to) those of the lottery pairs in one group, then if we also observe common ratio effects in choices in other groups of lotteries that have quite different slopes, they are most likely due to deviations from EUT, not due to behavioral noises.

In oder to estimate CPT, the lotteries in Harrison and Swarthout (2016) involve prizes both in the gain domain, the loss domain and the mixed domain. While some lotteries have prizes with positive probabilities only in the gain domain or only in the loss domain, many lottery pairs have them in both domains (referred to as mixed domain lotteries). Some lotteries lie on the boundaries of the MM triangles, which means one or two prizes are associated with zero probabilities; while other lotteries lie strictly in the interior, which means all three prizes represented in that triangle are associated with positive probabilities. Endowments are provided to cover possible losses when lotteries involve prizes in the loss domain with positive probabilities. Using the same 100 lottery pairs in the experiments conducted with undergraduate students, Harrison and Swarthout (2016) have two treatments on how they get the endowment: in some sessions subjects receive the endowments as "house money", while in other sessions subjects need to earn the endowment by taking some quizzes before the lottery tasks. In section 2.5, we use their data only from the "house money" sessions.

There are many important methodological points that an experimenter needs to consider in implementing Binary Lottery Tasks, one of which is how to determine a subject's payoff from her choices in all these tasks.[11] Some payoff protocols, such as the widely used Random Lottery Incentive Method in which the experimenter randomly chooses one among all lotteries that a subject has chosen to play out and to determine her payoff, are theoretically incompatible with CPT and have been empirically proved to introduce interdependence in subjects' decisions across tasks in Cox et al. (2015) and Harrison and Swarthout (2014). Notwithstanding these concerns, with some known preference parameters, we simulate subjects' choices in which the decision in one task is generated independently of other tasks, and then estimate the decision models under the assumption that a subject truthfully reveals her preference in each lottery pair without any considerations of the other pairs.

2.3 Models

Suppose we have N subjects and every subject chooses a preferred lottery in each of the T lottery pairs, therefore we will have T observations for each subject.[12] The two lotteries in each pair are labeled "left" or "right." A lottery is a probability distribution $\{p_1 \ \dots \ p_K\}$ on K prizes $x_1 \geq \ \dots \ \geq \ x_K$; some prizes are negative, which means subjects may lose money, therefore they will receive some house money endowment e in the beginning to cover possible losses. We

[11]Other considerations include whether to allow for the choice of indifference between the two lotteries, how to present the many lottery pairs to subjects, the order in which lottery pairs are presented, the interface for the lottery choices, whether to allow all subjects to have prior information on all the lottery pairs before making any choices.

[12]The lottery pairs could be the same or different for all subjects. We used simulated and observed choices over the same 100 lottery pairs for each subject.

consider two theories on how subjects evaluate the utilities they get from each lottery, EUT and CPT. We will first lay out the specifications of EUT and CPT, and then discuss the stochastic process of choices and derive the likelihood and posterior of the mixture model of these two models.

2.3.1 Expected Utility Theory (EUT)

In the evaluation of lottery prizes, assume individuals perfectly integrate the prizes with their endowments and behave as if they evaluate Constant Relative Risk Aversion (CRRA) utility functionals $u(e,x_k) = (e+x_k)^{(1-r)}/(1-r)$ for any $k = 1,...,K$. A lottery is evaluated by the weighted sum of utilities of prizes, with the weights being the objective probabilities associated with the prizes. Then, we have

$$EU = \sum_{k=1}^{K} p_k \cdot \left((e+x_k)^{(1-r)}/(1-r) \right)$$

The subject is assumed to select the lottery in a pair with the higher EU.

2.3.2 Cumulative Prospect Theory (CPT)

Kahneman and Tversky (1979) proposed the original version of prospect theory and later developed CPT in Tversky and Kahneman (1992) by incorporating the rank-dependent nature of probability weighting from Quiggin (1982). The key features of CPT are the sign-dependence in individuals' evaluations of lottery prizes and their perceptions of objective probabilities. The first deviation of CPT from EUT is that, instead of perfect integration between prizes and endowments, individuals evaluate prizes as gains or losses from some assumed reference point.[13] We assume that subjects use the received endowments e as reference points, and we use the family of CRRA

[13]While EUT typically assumes perfect asset integration, and CPT assumes no asset integration, Cox and Sadiraj (2006) generalized EUT by using utility functionals that allows for partial asset integration, which is the middle case between full and null asset integration.

utility functions as follows:

$$U(x_k) = {x_k}^{1-\alpha}/(1-\alpha) \text{ if } x_k \geq 0 \tag{2.1a}$$

$$U(x_k) = -\lambda \cdot (-x_k)^{1-\beta}/(1-\beta) \text{ if } x_k < 0 \tag{2.1b}$$

where λ is the loss aversion parameter, and α and β are the Arrow-Pratt measure of relative risk-aversion in the gain and loss domains, respectively.

Another deviation of CPT from EUT is that individuals sum over the utilities from prizes with the associated *decisions weights*, which may or may not be equal to the objective probabilities. We outline the calculation of decision weights using the approach provided by Wakker (2010). Define the *gain-* and *loss-ranks* of probabilities (ranking prizes from best to worst $x_1 \geq ... \geq x_K$):

$$r(p_k) = p_1 + p_2 + ... + p_{k-1} \qquad \text{Gain-Rank of } p_k,\ k > 1$$

$$l(p_k) = p_K + p_{K-1} + ... + p_{k+1} \qquad \text{Loss-Rank of } p_k,\ k < K$$

$$r(p_1) = 0 \qquad \text{Gain-Rank of } p_1$$

$$l(p_K) = 0 \qquad \text{Loss-Rank of } p_K$$

We allow for different probability weighting functions on the gain-ranks and loss-ranks associated with prizes in the gain and loss domains. Assume the following inverse-S functionals for the probability weighting functions:

$$\omega^+(p) = p^{\gamma+}/(p^{\gamma+} + (1-p)^{\gamma+})^{1/\gamma+}$$

$$\omega^-(p) = p^{\gamma-}/(p^{\gamma-} + (1-p)^{\gamma-})^{1/\gamma-}$$

where γ^+ and γ^- are the probability weighting coefficients in the gain and loss domains, respectively.

For a gain-domain lottery, which only has positive prizes that are associated with positive probabilities, we rank the prizes from best to worst such that $x_1 \geq ... \geq x_K \geq 0$. We apply the probability weighting function to the *gain-ranks* of the probabilities, and then find the decision weight for each prize at its rank:

$$
\begin{aligned}
&\pi(x_1) = \omega^+(p_1) \\
&\pi(x_2) = \omega^+(p_1 + p_2) - \omega^+(p_1) \\
&\quad ... \\
&\pi(x_K) = \omega^+(1) - \omega^+(p_1 + .. + p_{K-1})
\end{aligned}
$$

The cumulative prospective utility (CPU^+) from playing this gain-domain lottery is then

$$CPU^+ = \sum_{k=1}^{K} \pi(p_k)U(x_k) = \sum_{k=1}^{K} \pi(p_k) \cdot \left({x_k}^{1-\alpha}/(1-\alpha)\right)$$

For a loss-domain lottery, which only has negative prizes that are associated with positive probabilities, we rank the prizes from best to worst such that $0 \geq x_1 \geq ... \geq x_K$. We apply the probability weighting function to the *loss-ranks* of the probabilities, and then we calculate the decision weights as follows [14]

$$
\begin{aligned}
&\pi(x_K) = \omega^-(p_K) \\
&\pi(x_{K-1}) = \omega^-(p_K + p_{K-1}) - \omega^-(p_K) \\
&... \\
&\pi(x_1) = \omega^-(1) - \omega^-(p_K + .. + p_2)
\end{aligned}
$$

[14]Here we index the *ranking of prizes* the same as what we did in the gain-domain, and apply the probability weighting function to the loss-ranks of the probabilities starting with prize x_K. Equivalently, as in Harrison and Swarthout (2016), one can index the prizes from the worst to the best such that $x_1 \leq ... \leq x_K \leq 0$ and then apply the probability weighting functions starting with x_1. The two ways are mathematically identical – both are applying the probability weighting functions to loss-ranks of probabilities associated with losses.

The cumulative prospective utility (CPU^-) from playing this loss-domain lottery is then

$$CPU^- = \sum_1^K \pi(p_k)U(x_k) = \sum_1^K \pi(p_k)\cdot\big(-\lambda(-x_k)^{1-\beta}/(1-\beta)\big)$$

For a mixed-domain lottery, we rank the prizes with respect to 0 so that we have $x_1 \geq ... \geq x_k \geq 0 \geq x_{k+1} \geq ... \geq x_K$, and then parse the mixed-domain lottery into a gain-domain lottery $\{x_1, p_1; ...; x_k, p_k; 0, p_{k+1} + ... + p_K\}$ and a loss-domain lottery $\{0, p_1 + ... + p_k; x_{k+1}, p_{k+1}; ...; x_K, p_K\}$. We calculate the utilities from these two lotteries respectively, using the methods listed above:

The Gain-Domain Lottery:

$\pi^+(x_1) = \omega^+(p_1)$

$\pi^+(x_2) = \omega^+(p_1 + p_2) - \omega^+(p_1)$

...

$\pi^+(x_k) = \omega^+(p_1 + .. + p_k) - \omega^+(p_1 + .. + p_{k-1})$

$$\begin{aligned} CPU^+ &= \sum_{j=1}^{k} \pi^+(p_j)U(x_j) + \pi^+(0)U(0) \\ &= \sum_{j=1}^{k} \pi^+(p_j)\cdot\big({x_j}^{1-\alpha}/(1-\alpha)\big) \end{aligned}$$

The Loss-Domain Lottery:

$\pi^-(x_K) = \omega^-(p_K)$

$\pi^-(x_{K-1}) = \omega^-(p_K + p_{K-1}) - \omega^-(p_K)$

...

$\pi^-(x_{k+1}) = \omega^-(p_K + .. + p_{k+1}) - \omega^-(p_K + .. + p_{k+2})$

$$\begin{aligned} CPU^- &= \sum_{j=k+1}^{K} \pi^-(p_j)U(x_j) + \pi^-(0)U(0) \\ &= \sum_{j=k+1}^{K} \pi^-(p_j)\cdot\big(-\lambda(-x_j)^{1-\beta}/(1-\beta)\big) \end{aligned}$$

The utility of the mixed-domain lottery is the sum of the utilities from the parsed gain-domain lottery and the parsed loss-domain lottery:

$$CPU = CPU^+ + CPU^- = \sum_{j=1}^{k} \pi^+(p_j)\cdot\big({x_j}^{1-\alpha}/(1-\alpha)\big) + \sum_{j=k+1}^{K} \pi^-(p_j)\cdot\big(-\lambda(-x_j)^{1-\beta}/(1-\beta)\big)$$

While the idea of loss aversion is simply “losses loom larger than gains” (Kahneman and Tversky (1979; p.279)), the formal definition or measure of loss aversion are not agreed upon in the literature. Abdellaoui et al. (2007) summarized the popular definitions and estimation methods used in the literature to identify utility loss aversion,[15] and Harrison and Swarthout (2016; §2.E) provided a critical review. Following Wakker (2010), when evaluating lottery prizes within CPT, we differentiate between the role of *utility loss aversion* λ and the role of *basic utility* $u(x_k)$, and refer to $U(x_k)$ in (2.1a) and (2.1b) as *the overall utility*. Therefore, we have $U(x_k) = u(x_k)$ if $x_k \geq 0$ and $U(x_k) = \lambda u(x_k)$ if $x_k < 0$, where $u(x_k)$ is specified as

$$u(x_k) = {x_k}^{1-\alpha}/(1-\alpha) \text{ if } x_k \geq 0 \tag{2.2a}$$

$$u(x_k) = -(-x_k)^{1-\beta}/(1-\beta) \text{ if } x_k < 0 \tag{2.2b}$$

Based on this differentiation, we define utility loss aversion as $\lambda \equiv -U(-1)/U(1)$ following Tversky and Kahneman (1992).[16] Note this definition requires a scaling convention such that $u(1) = 1$ and $u(-1) = -1$, which was implicitly done in Tversky and Kahneman (1992) since they chose $u(x_k) = x_k^{\alpha}$ and $u(-x_k) = -(x_k)^{\beta}$. Under this definition, we use the same identification strategy as Harrison and Swarthout (2016; p.21): "[...] the empirical strategy is to evaluate estimates of α and β, and then infer λ by evaluating the implied utility functions at ± 1. Estimates of all three parameters are then used, along with estimated decision weights, to evaluate each lottery [...]"

2.3.3 Stochastically Linking Theory to Choices

Looking at a single pair of lotteries, we have outlined two distinct ways of evaluating each lottery in the pair, which we refer to as the *deterministic component* in the decision process.[17] Turning to the *stochastic component* of the decision process, assume subjects compare the two

[15] The name *utility loss aversion* is used to differentiate with the *probabilistic loss aversion*. The latter can be generated by different probability weighting functions in the gain and loss domains. For example, an individual exhibits probabilistic loss aversion if she has probability neutrality for gains and probability pessimism for losses.

[16] Examples of other definitions of loss aversion, as reviewed in Abdellaoui et al. (2007), include the ratio of the directional derivatives of overall utility at the reference point at the reference point $U'_{\rightarrow}(0)/U'_{\leftarrow}(0)$, and the ratio between overall utility loss and overall utility gain at a different unit of monetary or non-monetary payoffs $-U(-x)/U(x)$.

[17] Referring to the evaluations of lotteries as the "deterministic component" in the decision process reflects the view that preferences are fixed while choices are noisy. Another view is that preferences are inherently random (examples include the random utility theory discussed in Becker et al. (1963) and its generalization in Loomes and Sugden (1995)).

lotteries by taking the differences between the utilities of the two lotteries $\nabla V = (V_L - V_R)/\mu$, where V represents either EU or CPU as calculated above and μ represents a behavioral error parameter popularized by Hey and Orme (1994) and originally due to Fechner. A greater μ indicates a higher level of noises in decision making.

We assume a logistic link between the observed choices and the difference between the utilities of the two lotteries in a pair. Let $y = 1$ represent choosing the left lottery in each pair, then[18]

$$\text{Prob}(y=1) = G(\nabla V) = \frac{exp(\nabla V)}{1+exp(\nabla V)}$$

or

$$y \sim \text{Bernouli}\left(\frac{exp(\nabla V)}{1+exp(\nabla V)}\right)$$

2.3.4 Simulated Choices

We simulate 177 subjects' choices over the 100 lottery pairs using the decision processes described above, then we develop Bayesian algorithms to recover the preference parameters from these simulated choices and compare the estimates to the known, true parameters. In the simulation, instead of assuming all subjects are EUT *or* CPT, we assume they are a mixture of EUT *and* CPT types: 109 of the simulated subjects (approximately 60%) are characterized by EUT, while the others 68 (approximately 40%) are characterized by CPT. We use ρ to represent the mixture weight of EUT type in all 177 subjects. We assume subjects do not switch between the EUT and CPT types over all 100 lottery tasks, while providing a discussion on the estimation under the assumption that subjects do switch between the two types in Appendix E. The mixture weight,

[18]Other ways to define the link include the Probit link with $\text{Prob}(y=1) = \Phi(\nabla V)$, where $\Phi()$ is the cumulative distribution function of a standard normal distribution, and the ratio link which, instead of taking differences in the utility, assumes $y \sim \text{Bernouli}\left(\frac{V_L^{1/\mu}}{V_L^{1/\mu}+V_R^{1/\mu}}\right)$.

preferences parameters for EUT subjects and for CPT subjects are summarized as follows:

Percentage of EUT Subjects: $\rho = 0.6$

EUT CRRA: $r = 0.3$

EUT Noise: $\mu_{EUT} = 2$

CPT Gain-Domain CRRA: $\alpha = 0.1$

CPT Loss-Domain CRRA: $\beta = 0.5$

CPT Loss Aversion Parameter: $\lambda = 1.8$

CPT Gain-Domain Weighting Function: $\gamma^{+} = 0.8$

CPT Gain-Domain Weighting Function: $\gamma^{-} = 0.6$

CPT Noise: $\mu_{CPT} = 2$

2.4 Bayesian Models and Algorithms

In this section we derive the joint posterior of the parameters, describe two MCMC algorithms to sample from the posterior and report the results based on the simulated dataset described in the previous section.

Assume the behavioral noises are identically and independently distributed for all subjects and decisions. Collect all decisions of an arbitrary subject i in $y_i = (y_{i1}, \ldots, y_{i100})$, and collect all parameters in $\theta = (\rho, r, \mu_{EUT}, \alpha, \beta, \lambda, \gamma^{+}, \gamma^{-}, \mu_{CPT})$. Under the assumption that a subject uses one decision mode for all decisions, the likelihood of observing y_i is

$$p(y_i|\theta) = \rho \prod_{t=1}^{100} \left[G(\nabla EU_{it})^{y_{it}} \times (1 - G(\nabla EU_{it}))^{1-y_{it}} \right] + (1-\rho) \prod_{t=1}^{100} \left[G(\nabla CPU_{it})^{y_{it}} \times (1 - G(\nabla CPU_{it}))^{1-y_{it}} \right]$$

The likelihood of observing all subjects' decisions $y = \{y_1, \cdots, y_{177}\}$ is

$$p(y|\theta) = \prod_{i=1}^{177} p(y_i|\theta)$$

Assume parameters are independently distributed in the prior distribution, so the joint prior takes the form $p(\theta) = p(\rho)p(r)p(\mu_{EUT})p(\alpha)p(\beta)p(\gamma^+)p(\gamma^-)p(\mu_{CPT})$. Specify non-informative marginal prior for each parameter as

$$(\rho, 1-\rho) \sim \mathcal{D}(1,1),\ r \sim N(0,100), \tag{2.3a}$$

$$\alpha \sim N(0,100),\ \beta \sim N(0,100), \tag{2.3b}$$

$$\gamma^+ \sim N(0,100),\ \gamma^- \sim N(0,10), \tag{2.3c}$$

$$\mu_{EUT} \sim Uniform(0,100),\ \mu_{CPT} \sim Uniform(0,100) \tag{2.3d}$$

where $\mathcal{D}$ represents the Dirichlet distribution and N the Normal distribution.

Combining prior with likelihood we have posterior distribution $p(\theta|y) \propto p(y|\theta)p(\theta)$. Using MATLAB we write several MCMC algorithms to sample from this posterior. We describe these algorithms and present the results in the rest of this section.

2.4.1 The Metropolis Hastings (MH) Algorithms

This is a basic application of the MH algorithm developed by Metropolis et al. (1953). Use θ to represent a vector of arbitrary parameters. Then the m^{th} iteration of the algorithm is as follows:

(1) Propose candidate draws of θ using $\theta^{(c)} = \theta^{(m-1)} + \varepsilon$, $\varepsilon \sim N(0,c)$ where c is the proposal step constant. [19]

(2) Accept the candidate draws with probability $\max\{\frac{p(y|\theta^{(c)})p(\theta^{(c)})}{p(y|\theta^{(m-1)})p(\theta^{(m-1)})}, 1\}$

We developed three variations of this MH algorithms. In the first variation candidate draws of all parameters are proposed simultaneously and independently in step (1) and the posterior

[19] In the algorithm c is adjusted every 10 iterations to ensure the optimal acceptance rate as proposed by Gelman et al. (1996).

is evaluated at all candidate parameters to determine the acceptance probability; in the second variation parameters are partitioned into three blocks $\{\rho\}$, $\{r, \mu_{EUT}\}$ and $\{\alpha, \beta, \lambda, \gamma^{+}, \gamma^{-}, \mu_{CPT}\}$, candidate draws are proposed one block at a time, and are accepted based on the conditional posterior distribution of each block; in the third variation candidate draw is proposed for each individual parameter at a time, and is accepted based on the conditional posterior of each individual parameter. The second and third variations are essentially a combination of Gibbs sampler and MH algorithms.

After comparing the convergence quality of the three variations of MH algorithms, we find that the third variation has the best convergence and yields posterior samples with the lowest autocorrelations.[20] We tested the three variations with randomly drawn starting values of the parameters and find that the first two variations fail to converge when the starting values are very far from the true parameters and yield a very low posterior probability density, while the third variation almost always converges. Therefore, we only report the results of posterior samples using the third variation of this algorithm. Figure 2.2 displays the trace plot of the posterior sample for each parameter, which is the time-series plot with the values of parameter draws on the y-axis and the corresponding iterations on the x-axis. Figure 2.3 displays the histograms of posterior samples after burn-in, that is, we only use the posterior draws after the MCMC chains converge (burn in). Figure 2.4 displays the *lag-k* autocorrelations of each posterior sample from $k = 1$ up to $k = 40$, where the *lag-k* autocorrelation of an MCMC sample is the correlation coefficient between the two draws that are k iterations apart in the sample.

2.4.2 *The MH Algorithm with Data Augmentation for Allocations*

Let S_i represent the allocation of subject i to EUT $(S_i = 1)$ or CPT $(S_i = 0)$ type, and collect allocations for all subjects in $S = \{S_1, \cdots, S_{177}\}$. This algorithm samples the allocations from the

[20] Autocorrelation is an important measure of the efficiency of an MCMC algorithm: the lower the autocorrelation, the greater the amount of information contained in a given number of draws from the posterior (greater efficiency). For an arbitrary parameter θ and the posterior sample $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(n)}$, the *lag-k* autocorrelation is the correlation coefficient between $\theta^{(x)}$ and $\theta^{(x+k)}$.

conditional posterior distribution $p(S|y,\theta)$, and then uses MH algorithm to sample the parameters from $p(\theta|S,y)$.[21]

Collect relevant parameters for EUT and CPT types respectively in $\theta_{EUT} = (r, \mu_{EUT})$ and $\theta_{CPT} = (\alpha, \beta, \lambda, \gamma^+, \gamma^-, \mu_{CPT})$, then we have $\theta = (\rho, \theta_{EUT}, \theta_{CPT})$. First, establish for later use, the likelihood of observing a subject's decision y_i given that she is an EUT type is:

$$p(y_i|S_i = 1, \theta) = p(y_i|\theta_{EUT}) = \prod_{t=1}^{100} G(\nabla EU_{it})^{y_{it}} \cdot (1 - G(\nabla EU_{it}))^{1-y_{it}}$$

The likelihood of observing y_i given that she is a CPT type is

$$p(y_i|S_i = 0, \theta) = p(y_i|\theta_{CPT}) = \prod_{t=1}^{100} G(\nabla CPU_{it})^{y_{it}} \cdot (1 - G(\nabla CPU_{it}))^{1-y_{it}}$$

This algorithm starts with some initial value of $\theta^{(0)} = \{\rho^{(0)}, r^{(0)}, \mu_{EUT}^{(0)}, \alpha^{(0)}, \beta^{(0)}, \lambda^{(0)}, \gamma^{+(0)}, \gamma^{-(0)}, \mu_{EUT}^{(0)}\}$, and the m^{th} iteration is:

(1) Sample allocation $S_i^{(m)}$ from $p(S_i|y_i, \theta^{(m-1)})$:

$$\begin{aligned}
Pr(S_i = 1|y_i, \theta^{(m-1)}) &\propto p(y_i|S_i = 1, \theta^{(m-1)}) Pr(S_i = 1|\theta^{(m-1)}) \\
&= p(y_i|\theta_{EUT}^{(m-1)}) \rho^{(m-1)} \\
Pr(S_i = 0|y_i, \theta^{(m-1)}) &\propto p(y_i|S_i = 0, \theta^{(m-1)}) Pr(S_i = 0|\theta^{(m-1)}) \\
&= p(y_i|\theta_{CPT}^{(m-1)}) (1 - \rho^{(m-1)}) \\
S_i \sim \text{Bernoulli}&\left(\frac{Pr(S_i = 1|y_i, \theta^{(m-1)})}{Pr(S_i = 1|y_i, \theta^{(m-1)}) + Pr(S_i = 0|y_i, \theta^{(m-1)})} \right)
\end{aligned}$$

[21] This is an application of the algorithm introduced in Chapter 3 of Frühwirth-Schnatter (2006), and is similar to the algorithms used in Houser et al. (2004).

Collect the allocations for all the subjects in $S^{(m)} = \{S_1^{(m)}, \cdots, S_{177}^{(m)}\}$.[22]

(2) Sample parameters $\theta^{(m)}$ from $p(\theta^{(m)}|y, S^{(m)})$:

$$\begin{aligned}
p(y,S|\theta) &= \prod_{i=1}^{177} p(y_i|S_i,\theta)p(S_i|\theta) \\
&= \prod_{i:S_i=1}\Big(\rho \cdot p(y_i|\theta_{EUT})\Big) \times \prod_{i:S_i=0}\Big((1-\rho)\cdot p(y_i|\theta_{CPT})\Big) \\
&= \underbrace{\prod_{i:S_i=1} p(y_i|\theta_{EUT})}_{\text{Likelihood of EUT Subjects}} \times \underbrace{\prod_{i:S_i=0} p(y_i|\theta_{CPT})}_{\text{Likelihood of CPT Subjects}} \times \underbrace{\rho^{N_1}(1-\rho)^{N_0}}_{\text{Likelihood of Allocations}}
\end{aligned}$$

The superscripts for iteration are dropped for a cleaner look. N_1 is the number of subjects classified as EUT, and N_0 is the number of subjects classified as CPT from step (1). Under the assumption that parameters are independently distributed in their prior distribution, we have

$$\begin{aligned}
&p(\theta|y,S) \propto p(y,S|\theta)p(\theta) \\
&= \underbrace{\prod_{i:S_i=1} p(y_i|\theta_{EUT})\cdot p(\theta_{EUT})}_{\text{Posterior of EUT Subjects}} \times \underbrace{\prod_{i:S_i=0} p(y_i|\theta_{CPT})\cdot p(\theta_{CPT})}_{\text{Posterior of CPT Subjects}} \times \underbrace{\rho^{N_1}(1-\rho)^{N_0}\cdot p(\rho)}_{\text{Posterior of Allocations}}
\end{aligned}$$

Therefore, in step (2) of the m^{th} iteration, we can separately sample θ_{EUT}, θ_{CPT} and ρ from their respective posteriors. For the parameters within each type θ_{EUT} and θ_{CPT}, we use the MH algorithm with proposing functions:

$$\theta_{EUT}^{(c)} = \theta_{EUT}^{(m-1)} + \varepsilon_{EUT},\ \varepsilon_{EUT} \sim N(0, c_{EUT})$$

[22] Note the complete posterior distribution of S_i should be

$$\begin{aligned}
p(S_i|y_i,\theta) &= \frac{p(S_i,y_i,\theta)}{p(y_i,\theta)} = \frac{p(y_i|S_i,\theta)p(S_i,\theta)}{p(y_i,\theta)} \\
&= \frac{p(y_i|S_i,\theta)p(S_i|\theta)p(\theta)}{p(y_i,\theta)} \propto p(y_i|S_i,\theta)p(S_i|\theta)
\end{aligned}$$

$$\theta_{CPT}^{(c)} = \theta_{CPT}^{(m-1)} + \varepsilon_{CPT},\ \varepsilon_{CPT} \sim N(0, c_{CPT})$$

The step constants c_{EUT} and c_{CPT} are independently adjusted every 10 iterations to ensure good acceptance rates.

We can directly sample ρ from its posterior distribution, since we use a Dirichlet prior for this parameter which is conjugate to the binomial likelihood of allocations $p(S|\rho) = \rho^{N_1}(1-\rho)^{N_0}$. Specify the Dirichlet prior as $\left(\rho, 1-\rho\right) \sim \mathcal{D}(e_{0,1}, e_{0,2})$, hence the posterior will be $\left(\rho, 1-\rho \middle| S\right) \sim \mathcal{D}\left(e_1(S), e_2(S)\right)$ where

$$e_1(S) = e_{0,1} + N_1,\ e_2(S) = e_{0,2} + N_0$$

The mixture model can be treated as an missing data problem, where the missing data is the preference type of each subject. This algorithm samples the missing allocation of each subject to EUT or CPT in step (1) conditional on the last draws of preference parameters and mixture probability, a technique that is commonly referred to as data augmentation. The benefit of data augmentation on the allocations (label augmentation) is that after categorizing each subject as either EUT or CPT type, the joint posterior distribution can be separated into three independent marginal posterior distributions as shown in step (2), which tends to make the MCMC chain converges faster and the algorithm more efficient. In addition, the marginal posterior of mixture probability ρ has a closed form after we augment the allocations when we use the conjugate Dirichlet prior, which makes the algorithms more efficient as well. Apart from the considerations of efficiency, another minor advantage of this algorithm is that the type of each subject as EUT or CPT can be acquired as a by product from step (1) of the estimation.

We also have two variants on the algorithm introduced above. In one variation the EUT or CPT parameters are updated in one MH step, while in the second variation each parameter is updated individually from its conditional posterior distribution. Between these two algorithms, we also find the individually updated variation has better convergence quality and the samples have lower autocorrelation. The trace plots of the samples from the second variant of this algorithm are graphed in Figure 2.5, the histograms after burn-in in are graphed in Figure 2.6, and the autocorrelation plot in Figure 2.7.

Comparing the algorithms with or without label augmentations, the convergence quality seem to be equally good. The sample means and standard deviations are very similar between the two types of algorithms. Table 2.1 reports the means, standard deviations, 5% quantile and 95% quantile of the posterior samples from both algorithms, and lists true values of these parameters in the first column as reference for comparison. The 5% and 95% quantiles can be thought of as the lower and upper bound of the 90% confidence intervals. Observing that the true values indeed fall within the 90% confidence interval, we conclude that we are successful in recovering the true parameters

Table 2.1: Comparison between True Parameters and Bayesian Estimation Results

Paramters	True Values	Metropolis Hasting Without Label Augmentation				Metropolis Hasting With Label Augmentation			
		Mean	S.D.	5% Quantile	95% Quantile	Mean	S.D.	5% Quantile	95% Quantile
ρ	**0.6**	0.615	0.035	0.557	0.672	0.614	0.036	0.554	0.673
r	**0.3**	0.310	0.009	0.295	0.324	0.309	0.009	0.294	0.324
μ_{EUT}	**2.0**	1.978	0.072	1.860	2.100	1.983	0.074	1.865	2.105
α	**0.1**	0.101	0.011	0.082	0.119	0.100	0.012	0.081	0.119
β	**0.5**	0.515	0.028	0.468	0.561	0.518	0.028	0.472	0.564
λ	**1.8**	1.860	0.114	1.679	2.056	1.872	0.114	1.698	2.078
γ^+	**0.8**	0.806	0.011	0.789	0.825	0.805	0.011	0.787	0.823
γ^-	**0.6**	0.595	0.018	0.563	0.624	0.594	0.018	0.563	0.624
μ_{CPT}	**2.0**	2.021	0.106	1.858	2.211	2.032	0.107	1.860	2.214

2.5 Applications to Observed Data

We now report results based on the observed subject choices rather than the simulated choices. The observed data refers to the data collected in the experiments reported by Harrison and Swarthout (2016), in which 177 recruited subjects make decisions over the same 100 lottery pairs as in the simulation data. After all decisions are made, one of the decisions was randomly chosen to determine subjects' payoffs from the experiment. We apply the algorithms introduced in Section 2.4 under the assumption that subjects do not switch between being an EUT or CPT type and with the same non-informative priors as in Section 2.4. The trace plots, histograms and autocorrelation plots are reported in Figures 2.9, 2.10 and 2.11, respectively. Summary statistics are reported in Table 2.2.

We find that 56.5% subjects can be categorized as the EUT type. Among the EUT types, we find moderate risk aversion with $r = 0.640$. Turning to the CPT types, we have graphed the estimated utility and probability weighting functions in Figure 2.8. We find that subjects tend to overestimate probabilities associated with extreme outcomes in both gain and loss domains, but the overweighting is more prominent in the gain domain than in the loss domain. Turn to the basic utility $u(x)$: in the gain domain, $u(x)$ is practically linear, and in the loss domain, $u(x)$ is slightly convex. Although the absence of utility loss aversion can be rejected at the 90% confidence level since $\lambda = 1$ is below the 5% quantile of the posterior sample, the average of the posterior sample for λ is very close to 1.

However, considering that we only estimated a mixture model of EUT and CPT with one set of functionals, and that we assumed away individual heterogeneity within each preference type, the results reported here should not be interpreted as our final conclusion on what best characterize the data sample.

Table 2.2: Bayesian Estimation Results Using Observed Subject Choices

Paramters	Mean	S.D.	5% Quantile	95% Quantile
ρ	0.565	0.040	0.500	0.631
r	0.640	0.008	0.627	0.654
μ_{EUT}	1.206	0.035	1.149	1.265
α	0.037	0.029	-0.010	0.083
β	0.138	0.031	0.086	0.190
λ	1.118	0.043	1.055	1.193
γ^{+}	0.690	0.017	0.661	0.719
γ^{-}	0.903	0.056	0.820	1.001
μ_{CPT}	6.068	0.625	5.141	7.155

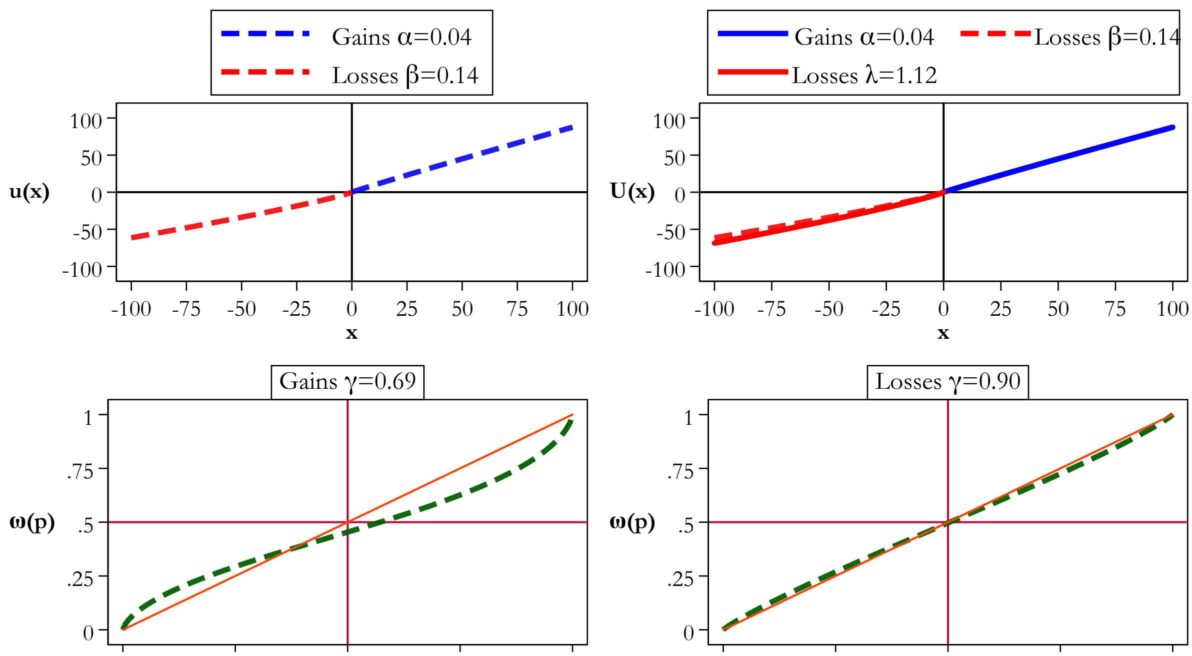


Figure 2.8: CPT Estimates Using Observed Subject Choices

2.6 Concluding Remarks and Future Plans

As the ultimate goal of the planned research project, we apply Bayesian econometric methods to address potential problems due to small sample size in the individual level analysis of subjects' risk preferences. As the first step, we apply Bayesian Econometric methods to the structural estimation of a mixture model of individual preferences under risks using pooled data in the current paper. We develop several MCMC algorithms to sample from the posterior distributions of the mixture model parameters and compare their efficiencies. In order to evaluate the algorithms, we first apply them to simulated data and find that two variants of the algorithms can successfully and efficiently recover the known, true parameters that are used to generate the stimulated data. We pay special attention to the technique of label augmentation and compare the algorithms with and without this technique. Although convergences are equally good between the two types of algorithms, the algorithms with label augmentation are still appealing as the type of each simulated subject can be acquired as a byproduct of the estimation.

We also apply the algorithms to the analysis of the observed data collected in the experiments of Harrison and Swarthout (2016). The results indicate that 56.5% subjects can be categorized as consistent with EUT. Similar to the results in the literature, the EUT type subjects exhibit moderate risk aversion, and the CPT type subjects overweight probabilities of extreme outcomes (the inverse-S shaped probability weighting functions in both gain- and loss-domains). In contrast to the evidences in the literature, the CPT type subjects exhibit very mild utility loss aversion. However, we are aware of the limitations of these results as we did not consider individual heterogeneity on preference parameters within each type. Also, we only estimated a mixture of EUT and CPT and did not explore mixtures of other preference models, and therefore have not undertaken a thorough model comparison across different models as in Harrison and Swarthout (2016).

For the next steps, we are going to expand our analysis on the observed data. As outlined in the introduction of the project, we also plan to address the issue of within-type heterogeneity and recover preference parameters at the individual level. We will do so by estimating the mixture model using a Hierarchical Bayesian modeling approach. We will refrain from the simplifying assumption that subjects of the same preference type have the same parameters, and assume they have different parameters which follow some probability distribution. If we estimate for each individual independently, without imposing a distribution over the parameters, due to the small sample of choice data we get for each subject and the high dimensionality of the parameter space, we expect to get very noisy estimates or even fail to get any estimates. The hierarchical structure introduces dependence of preference parameters across subjects, which we hope will help reduce the noise in the estimates and improve the quality of inference.

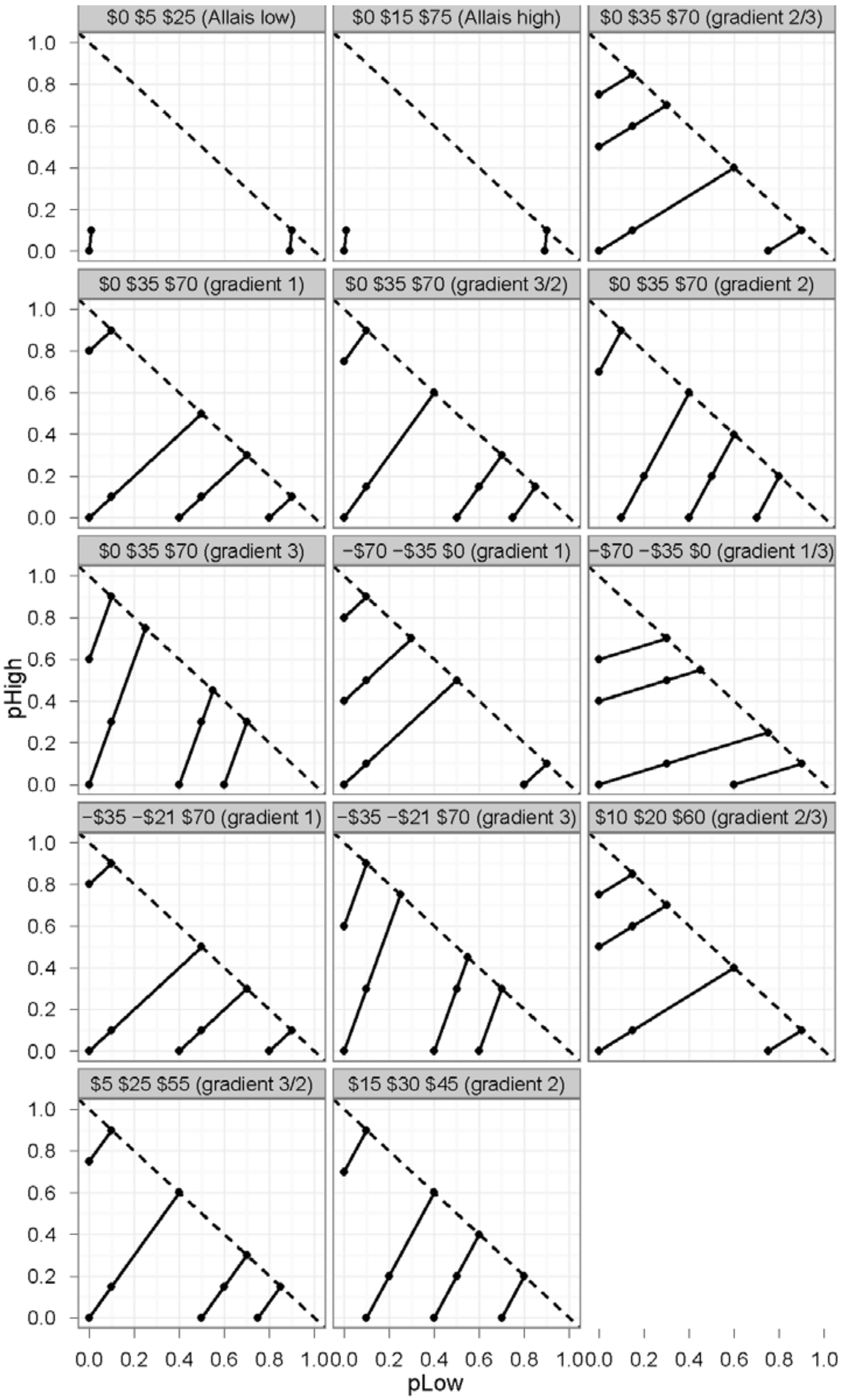
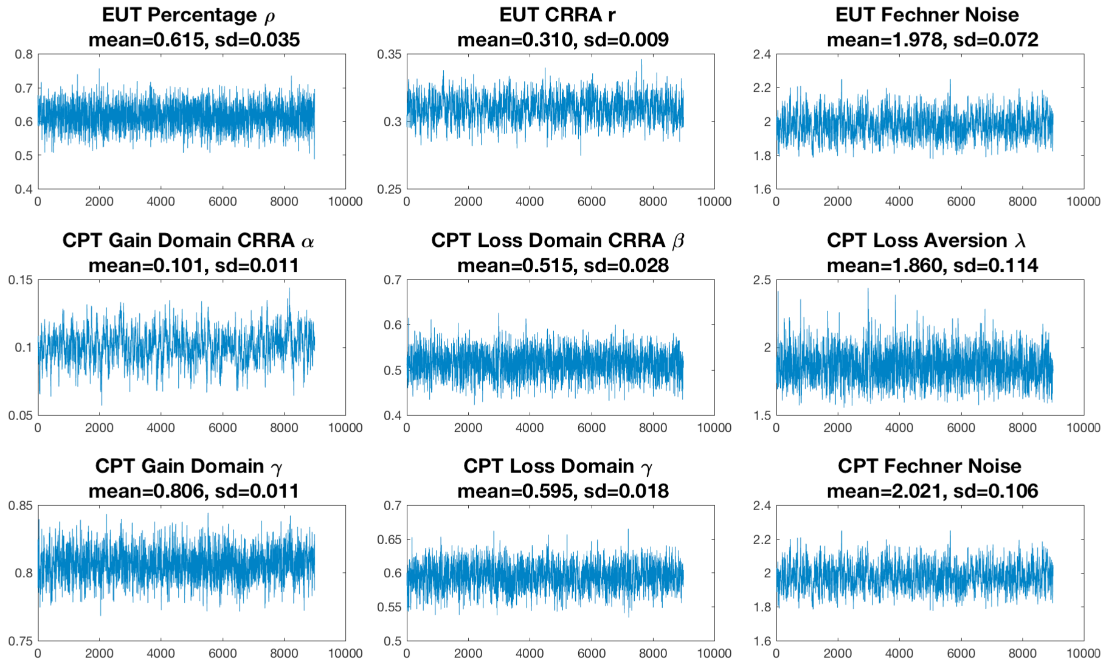


Figure 2.1: Lottery Pairs in Marschak-Machina Triangles



[1] The mean and standard deviation for each chain (after burn in) are noted above each graph.
[2] In Figures 2.2, 2.3 and 2.4: Assume subjects do not switch between EUT and CPT in data simulation and in the estimation.

Figure 2.2: Trace Plots of MH Samples Based on Simulated Observations

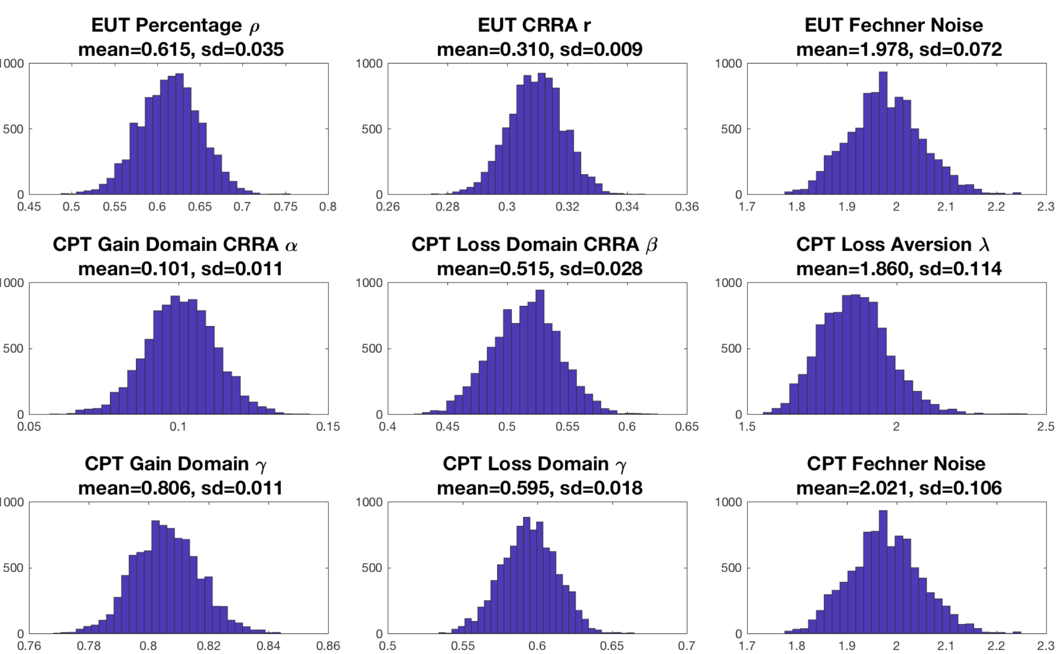


Figure 2.3: Histograms of MH Samples Based on Simulated Observations

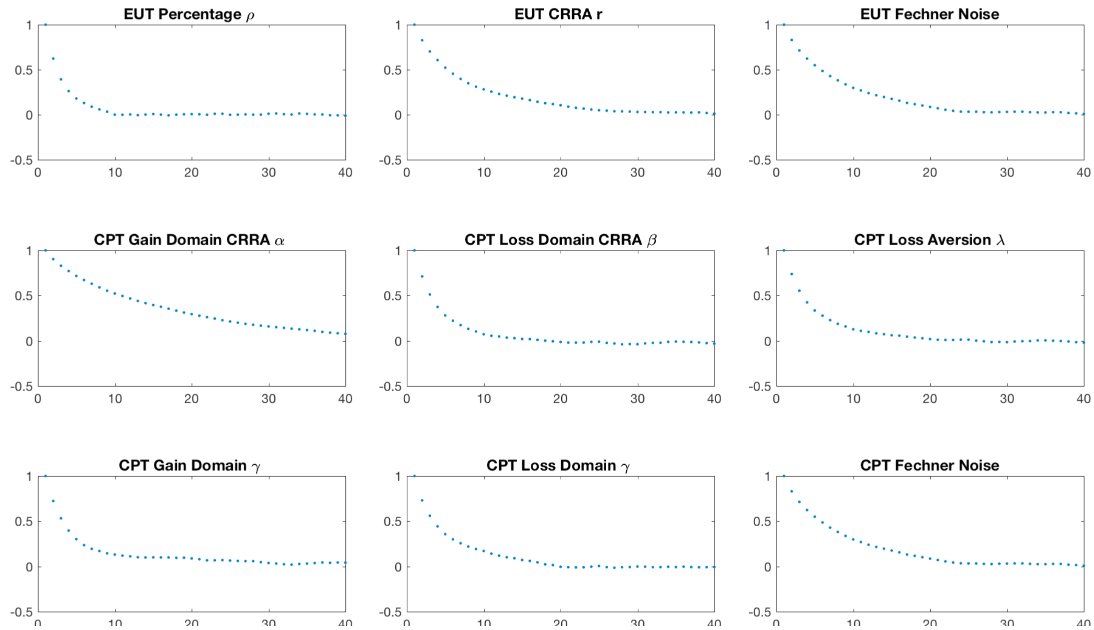
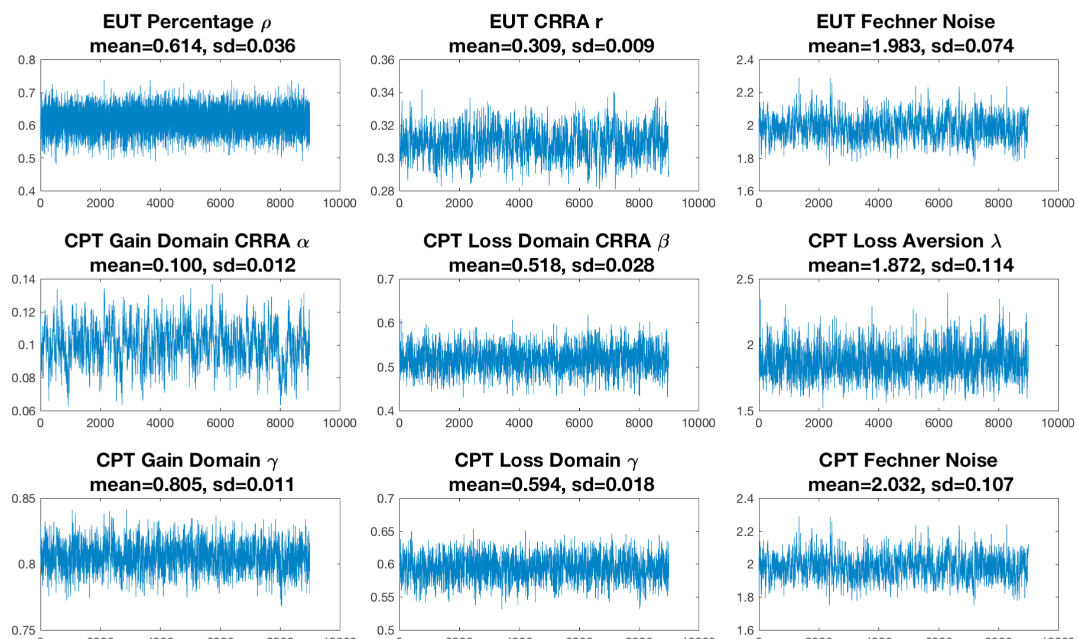


Figure 2.4: Autocorrelation of MH Sample Based on Simulated Observations



[1] The mean and standard deviation for each chain (after burn in) are noted above each graph.
[2] In Figures 2.5, 2.6 and 2.7: Assume subjects do not switch between EUT and CPT in data simulation and in the estimation.

Figure 2.5: Trace Plots of MH Samples with Augmentations for Labels

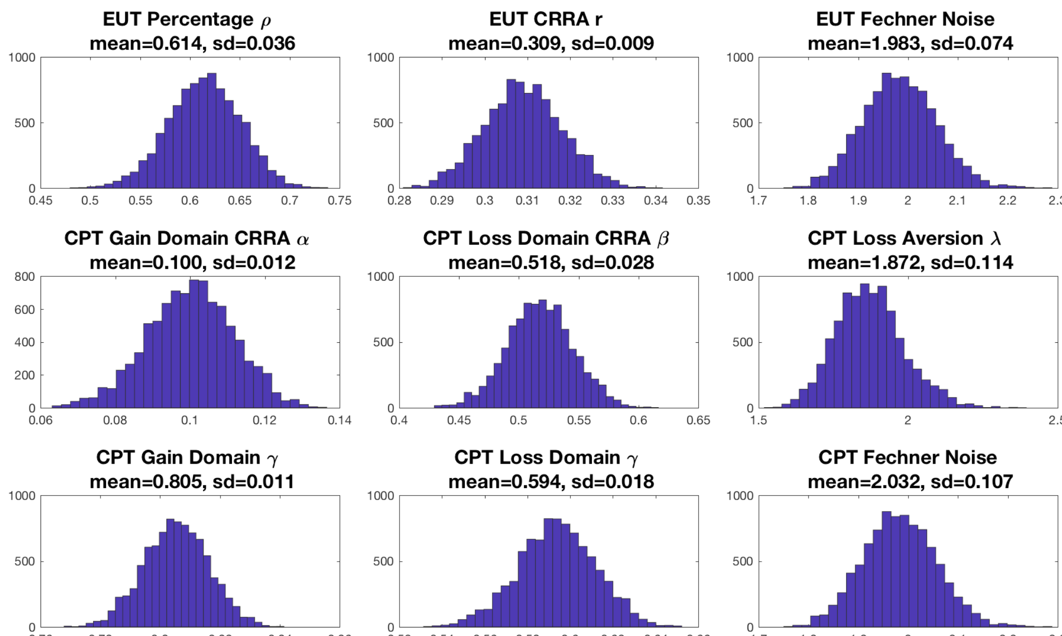


Figure 2.6: Histograms of MH Samples with Augmentations for Labels

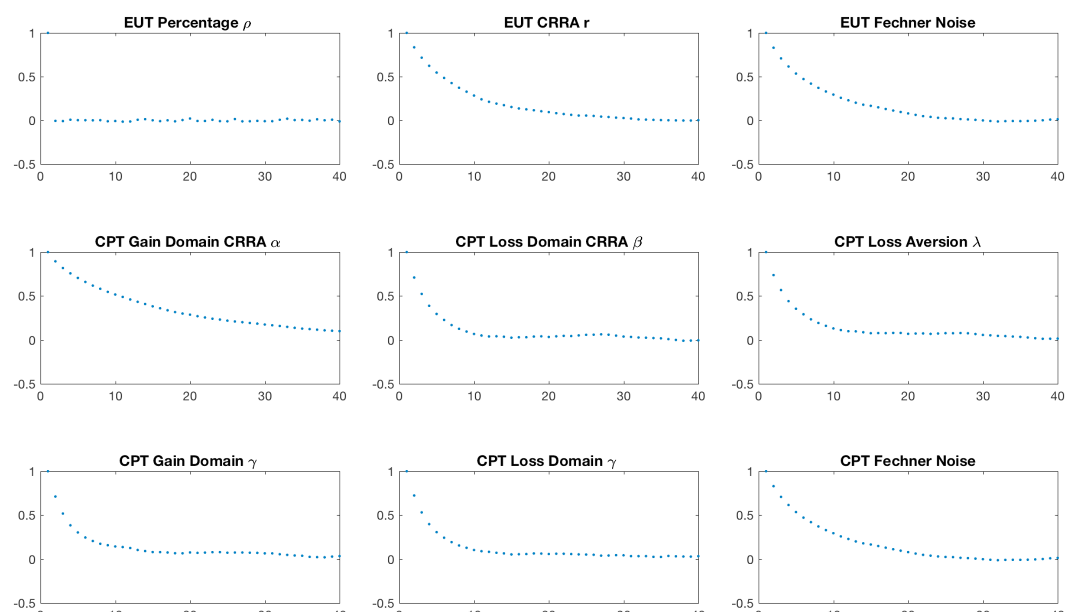
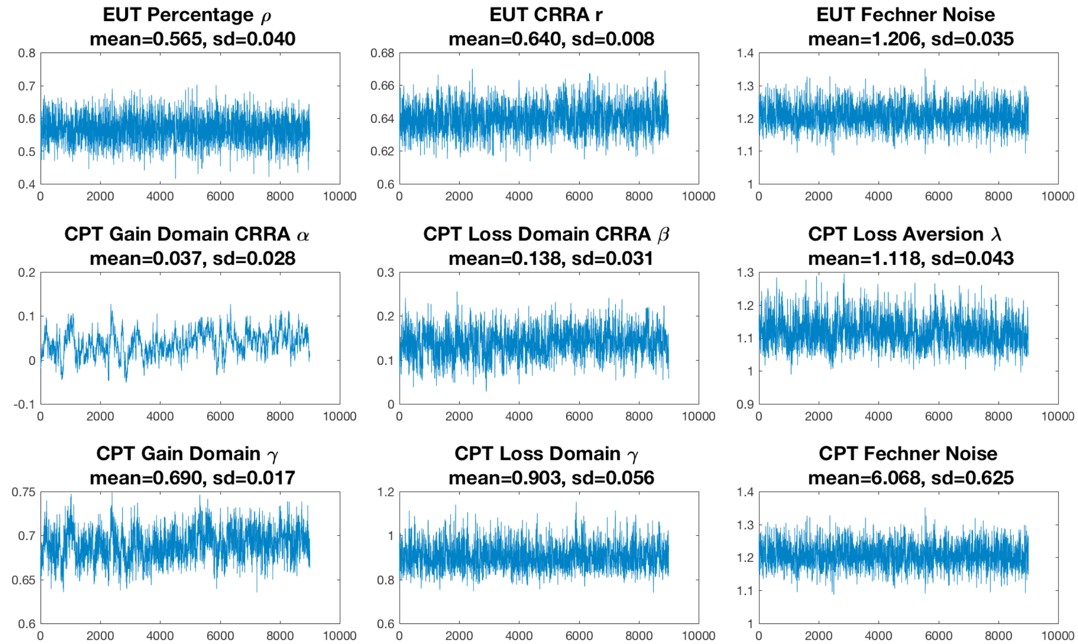


Figure 2.7: Autocorrelation of MH Samples with Augmentations for Labels



[1] In Figures 2.9, 2.10 and 2.11: Observed subjects' choices from Harrison and Swarthout (2016).
[2] In Figures 2.9, 2.10 and 2.11: Assume subjects do not switch between EUT and CPT.
[3] The mean and standard deviation for each sample are noted above each panel.

Figure 2.9: Trace Plots of Metropolis Hastings Samples (Observed Data)

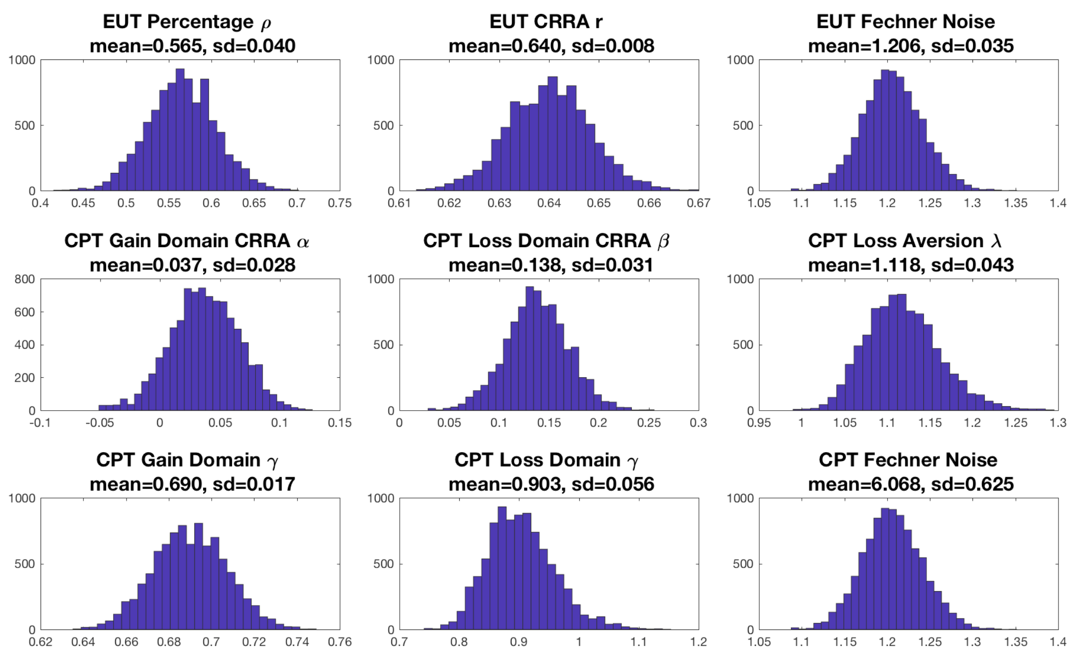


Figure 2.10: Histograms of Metropolis Hastings Samples (Observed Data)

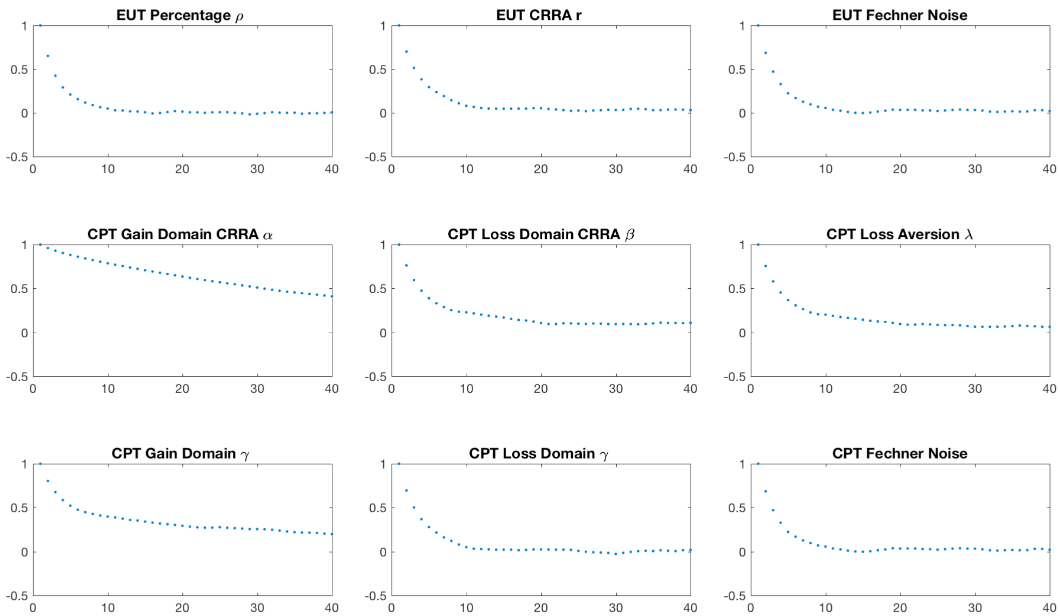


Figure 2.11: Autocorrelation Plots of Metropolis Hastings Samples (Observed Data)

CHAPTER 3

DISCHARGE DECISION-MAKING AFTER COMPLEX SURGERY: SURGEON BEHAVIORS COMPARED TO PREDICTIVE MODELING TO REDUCE SURGICAL READMISSIONS[1]

3.1 Introduction

The Centers for Medicare and Medicaid Services (CMS) have placed increased scrutiny on hospital readmissions. (Jencks et al. (2009); Dawes et al. (2014); Centers for Medicare & Medicaid Services (2007)) As mandated by the Patient Protection and Affordable Care Act, CMS has begun adjusting hospital payments through the Hospital Readmissions Reduction Program according to hospitals' rate of "excess" vs. "expected" Medicare readmissions for pneumonia, acute myocardial infarction, and heart failure with a future planned expansion into surgical patients.(Dawes et al. (2014); of the Patient Protection and Act (2010); Horwitz et al. (2011); Tsai et al. (2013); Joynt and Jha (2013)) Previous estimates suggest that even a small reduction of 5% in readmission rates could prevent over 2,000 inpatient hospitalizations with Medicare cost savings of $31 million. (Lawson et al. (2013))

[1]Coauthored with Ira L. Leeds , Vjollca Sadiraj, James C. Cox, Timothy M. Pawlik, Kurt E. Schnier and John F. Sweeney, published in The American Journal of Surgery 213 (2017) 112-119. Permission to reproduce this paper is granted by ©2016 Elsevier Inc. The published paper can be found at http://dx.doi.org/10.1016/j.amjsurg.2016.03.010.

One of the surgeon's most challenging clinical decisions is balancing the need to promptly discharge patients versus a clinical and financially incentivized goal of reducing readmissions. (Lawson et al. (2013); Bueno et al. (2010); Kaboli et al. (2012))

Balancing countervailing needs has often been addressed through the use of risk-based modeling and decision-support tools. The financial implications of readmissions have also led to many scientific inquiries into risk-adjusted predictions for readmission. A recent systematic review found 26 unique models of readmission employing a variety of data sources and types of inpatient populations.(Kansagara et al. (2011)) An ongoing limitation of these prediction tools has been the decreasing statistical discrimination of models when broadening patient populations to include surgical patients, especially those undergoing a wide variety of procedures. (Kansagara et al. (2011); Kassin et al. (2012); Glance et al. (2014); Morris et al. (2014); Merkow et al. (2015))

We believe that improving discharge decision-making via evidence-based decision-support tools will lower readmissions while maintaining or decreasing LOS. This approach requires two central elements: (a) statistical identification of variables that discriminate between likelihood of discharge and likelihood of subsequent readmission; and (b) development of decision-support software that can aid discharge decision-making by effectively operationalizing this risk-adjusted understanding of readmission into the clinical provider's daily work.

We sought to develop a data-driven predictive model for surgical readmission to identify the association between clinical information used for discharge decision-making and patients' subsequent risk of readmission. Retrospective, large-data analysis of a prospectively collected clinical data warehouse was used in a time-to-event model to identify criteria that (statistically) explain timing of inpatient postoperative discharge. Subsequent development of a prediction model of readmission with validation helps identify dissonant criteria across postoperative discharges and readmissions. Specifically, we researched possible discordance between intrinsic human behavior and optimized modeling with the assumption that such discordance could interfere with future uptake of decision-support tools. In particular, we wanted to identify differences in how surgeons behave in practice and how a predictive model of readmission might improve discharge decision-making.

3.2 Materials and methods

3.2.1 Patient population

De-identified patient data from all patients undergoing inpatient general (including gastrointestinal, endocrine, skin and soft tissue) and vascular surgical procedures between 2009 and 2014 were obtained from the academic medical center's clinical data warehouse. Both elective and emergency cases were included and controlled for in the models described below. Patients who were dead at discharge were excluded. This dataset included all electronically collected information during the patient's admission including demographic information, procedures performed, medications administered, laboratory test results, diagnostic imaging, and nursing documentation. Readmissions were captured by repeat encounters within 30 days of index admission. Outside hospital encounters that did not result in a transfer back to the index hospital could not be obtained.

3.2.2 Designing discharge and readmission models with validation

Time-varying and fixed data for all patients were analyzed using a time-to-event regression model to identify significant time-point predictors of discharge on a given hospital day and a logit regression model to identify significant predictors of readmission. With the exception of the dummy variables, standardized values of all independent variables were used in both models. Both models included 23 procedural grouping variables (e.g., colectomy, hepatectomy, ventral hernia repair) to control for the type of procedure performed. In addition, we created a dummy variable ("Pre-Optimize") to control for patients who were admitted for a surgical procedure with the procedure delayed beyond the initial day of admission.

A Cox survival model was used for time-point (i.e., daily) discharge predictions allowing for different baseline hazards across procedures. Time-varying variables were grouped for analysis by hospital day. Variables reported more than once daily (up to 3) were averaged. Patients with a missing variable on the hospital day examined had the last known observed value of that variable carried over (i.e., step imputation). If a variable was never recorded for the entire hospital stay, the normalized value (i.e., mean of the upper and lower limit) of that variable within the population was used for all hospital days. Using other methods of imputation did not meaningfully change the predictive factors of the model. A logit model with procedure-fixed effects was used to model readmissions using data from the day prior to discharge; a time-to-event specification for readmissions was not possible without time-point data following discharge. All explanatory variables were selected for using stepwise Akaike information criterion thresholds, which also accounted for Type I multiple testing error. (Burnham and Anderson (2003))

Both models were validated via a series of in-sample and out-of-sample tests using bootstrapped, partitioned patient data and C statistic test for discrimination. Iterations were conducted with a 90% in-sample and 10% out-of-sample partition, 70% in-sample and 30% out-of-sample partition, and a 50% in-sample and 50% out-of-sample partition. The normalized regression estimated coefficients of the empirical discharge model and the readmission predictive model were directly compared. All statistical analyses and modeling were performed using Stata® version 14.0 (StataCorp, College Station, TX).

Both methodologies were reviewed and approved by the Emory University and Georgia State University Institutional Review Boards.

3.3 Results

A total of 20,970 patients were identified from the institution's clinical data warehouse representing a wide range of surgical procedures. The median age of the patient population was 54 (range 13-96); 38.8% were male; 57.7% were white and 33.0% were black. Patients had a median length of stay after surgery of 2 days, and the distribution was skewed toward patients with prolonged lengths of stay (mean = 5.5 days, IQR = 1-6 days). Common comorbidities such as cancer (11.2%), hypertension (39.5%), and diabetes (16.1%) were frequently observed. The majority of operations (69.5%; 14,570) were gastrointestinal in nature. The 30-day readmission rate was 7.5% (1,565 patients), which is comparable to previously reported rates in general surgery patients. (Dawes et al. (2014); Lawson et al. (2013); Morris et al. (2014); Lucas et al. (2013)) Demographics, comorbidities, and surgical procedures of the patient population are fully described in Table 3.1.

Table 3.1: **Study population summary statistics.** 20,970 patients' daily clinical observations were extracted from an institutional data warehouse for all inpatient general and vascular surgery procedures from 2009 to 2014.

Age, median	54
Age, range	13-96
LOS, median (days)	2
LOS, mean (days)	5.53
	n (%)
Sex	
Male	8,143 (38.8%)
Female	12,827 (61.2%)
Race	
White	12,101 (57.7%)
Black	6,913 (33.0%)
Other	1,956 (9.3%)
Comorbidities	
Diabetes	3,376 (16.1%)
Cancer	2,354 (11.2%)
Heart disease	1,203 (5.74%)
Hypertension	8,270 (39.5%)
Procedure Category	
Gastrointestinal	14,570 (69.5%)
Endocrine	3,097 (16.8%)
Skin and Soft Tissue	2,668 (12.7%)
Ortho	426 (2.03%)
Thoracic	87 (0.4%)
Vascular	122 (0.6%)
30-day readmission	1,565 (7.47%)
Total patients	**20,970**
Total patient-days of observations	**115,976**

3.3.1 *Predictive modeling*

For each of the 23 procedure groups, 41 observed measures were routinely collected and stored in the clinical data warehouse (64 total collected variables). All of these were obtained for each of the 20,970 patients described above totaling 115,976 patient-days, including 1565 readmitted patients with 20,560 patient-days. These datasets were then used for discharge and readmission regression analyses. All 41 observed variables were included as regressors in both models after demonstrating the lack of multi-collinearity (mean variance inflation factor = 1.51; range 1.01e3.52) and survival via stepwise selection. 34 variables were found to be predictive (at conventional level of significance) of the discharge decision-making. Post hoc analysis suggested that difficult post-operative, in-hospital recoveries were indicative of increased readmission risk, so length of stay was added to the readmission model. 30 of the 42 variables were found to be significantly predictive of readmissions. The complete regression results for both models are reported in Table 3.2 and Table 3.4.

3.3.2 *Model validation*

Both models were tested with serial in-sample and out-of-sample bootstrapped, partitioned iterations. The discharge model employed 200 total iterations and the readmission model performed 300 total iterations with equal number of iterations conducted with 90% in-sample and 10% out-of-sample partitions, 70%/ 30%, and 50%/50%, respectively (50/50 for readmission model only due to regression methodology). The mean in-sample and out-of-sample C statistics were found to be 0.82 and 0.80 for the readmission model whereas for the discharge model the in-sample as well as out-of-sample mean C-statistic was 0.79.

Table 3.2: **Normalized effect of 42 observed clinical variables on discharge and readmission (n = 20,970 patients with 115,976 daily observed clinical measures; 1565 readmitted patients with 20,560 daily observed clinical measures).** Column A reports the normalized regression coefficient for each clinical variable for a time-to-event model analysis predicting daily likelihood of discharge based on actual discharges of the patient population. A positive value indicates the presence or increased magnitude of the variable increases the likelihood of discharge, and vice versa. Column B reports the normalized coefficient for each clinical variable for a logit regression model predicting likelihood of 30-day readmission as predicted on the day of discharge. A positive value indicates the presence or increased magnitude of the variables increases the likelihood of readmission. Confidence intervals of the coefficient estimates are reported to aid in the interpretation of conventional p-values for large sample sizes. Each patient was also controlled for the procedure performed (Table 3.4).

	(A)				(B)			
Clinical Variable	Discharge Regression Estimated Coefficient	95% CI		p	Readmission Regression Estimated	95% CI		p
Ascites	-0.452	-0.546	-0.358	0.000 ***	0.193	-0.019	0.404	0.074 *
CHF	-0.193	-0.258	-0.128	0.000 ***	0.275	0.123	0.427	0.000 ***
DM	-0.038	-0.071	-0.004	0.026 **	0.254	0.149	0.358	0.000 ***
CANCER	-0.167	-0.366	0.031	0.099 *	0.13	-0.015	0.276	0.080 *
HTN	-0.052	-0.086	-0.019	0.002 ***	0.144	0.027	0.261	0.016 **
Emer Admit	-0.075	-0.220	0.070	0.311	0.08	-0.175	0.336	0.538
Reg Admit	-0.098	-0.182	-0.013	0.023 **	0.074	-0.204	0.353	0.602
Male Gender	-0.045	-0.096	0.005	0.080 *	0.061	-0.004	0.125	0.066 *
PtAge	-0.043	-0.072	-0.013	0.005 ***	0.004	-0.080	0.089	0.921
WhiteRace	-0.046	-0.069	-0.022	0.000 ***	0.279	0.087	0.472	0.005 ***
BlackRace	-0.114	-0.164	-0.064	0.000 ***	0.366	0.133	0.599	0.002 ***
IVDU	-0.481	-0.689	-0.273	0.000 ***	0.251	0.158	0.343	0.000 ***
NPO	-2.069	-2.908	-1.229	0.000 ***	-0.331	-0.875	0.213	0.233
Albumin	-0.021	-0.031	-0.012	0.000 ***	-0.04	-0.062	-0.018	0.000 ***
AlkPhos	0.017	-0.023	0.056	0.403	0.075	0.029	0.121	0.001 ***
AST/ALT	-0.039	-0.063	-0.016	0.001 ***	-0.186	-0.337	-0.035	0.016 **
BUN	-0.170	-0.228	-0.113	0.000 ***	0.086	-0.012	0.184	0.084 *
Calcium	0.182	0.108	0.256	0.000 ***	-0.178	-0.236	-0.120	0.000 ***
Bicarb	-0.002	-0.028	0.024	0.865	-0.01	-0.056	0.036	0.665
Cr	0.014	-0.033	0.062	0.552	-0.036	-0.105	0.034	0.313
HCT	0.070	0.018	0.121	0.008 ***	-0.138	-0.179	-0.097	0.000 ***
INR	0.008	-0.010	0.026	0.387	-0.015	-0.049	0.020	0.408
K+	-0.035	-0.072	0.003	0.072 *	0.13	0.068	0.191	0.000 ***
Sodium	-0.061	-0.121	-0.002	0.043 **	-0.08	-0.130	-0.031	0.002 ***
PLT	0.075	0.050	0.100	0.000 ***	-0.042	-0.084	0.000	0.048 **
PTT	-0.016	-0.029	-0.004	0.009 ***	0.056	0.028	0.084	0.000 ***
TBili	-0.109	-0.151	-0.067	0.000 ***	0.065	-0.006	0.137	0.073 *
WBC	-0.190	-0.233	-0.147	0.000 ***	0.112	0.020	0.203	0.017 **
Katz Score	0.314	0.271	0.356	0.000 ***	-0.062	-0.098	-0.026	0.001 ***
Stool+	0.008	0.005	0.011	0.000 ***	-0.153	-0.292	-0.013	0.032 **
BMI	0.035	0.020	0.051	0.000 ***	-0.01	-0.086	0.066	0.792
HR	-0.166	-0.203	-0.130	0.000 ***	0.108	0.030	0.187	0.007 ***
O2Sat	-0.026	-0.039	-0.013	0.000 ***	-0.026	-0.060	0.007	0.127
RR	-0.003	-0.031	0.025	0.819	0.11	0.028	0.192	0.009 ***
DBP	0.017	-0.008	0.042	0.179	0.065	0.026	0.104	0.001 ***
SBP	-0.024	-0.040	-0.008	0.003 ***	-0.084	-0.133	-0.035	0.001 ***
TEMP	0.003	-0.030	0.036	0.847	0.048	0.005	0.092	0.031 **

PainScore	-0.068	-0.089	-0.047	0.000 ***	0.158	0.130	0.185	0.000 ***
DxImg	-0.650	-1.282	-0.018	0.044 **	0.175	-0.020	0.369	0.078 *
Complexity (RVU	-0.198	-0.364	-0.032	0.020 **	0.069	-0.027	0.165	0.161
UnivHosp vs Oth	-0.469	-0.567	-0.370	0.000 ***	0.372	0.145	0.598	0.001 ***
HighRisk	-0.856	-0.986	-0.726	0.000 ***	-0.175	-0.311	-0.039	0.011 **
LOS					-0.015	-0.023	-0.006	0.001 ***
LnLOS					0.659	0.504	0.814	0.000 ***

***p<0.01, **p<0.05, *p<0.1.

Abbreviations: Ascites – history of ascites, CHF – history of congestive heart failure, DM – history of diabetes, CANCER – history of cancer, HTN – history of hypertension, Emer Admit – unplanned emergency admission, Reg Admit – elective admission, PtAge – patient's age, IVDU - intravenous drug use, NPO – oral diet prohibited, Alk Phos – serum alkaline phosphatase, AST/ALT – serum liver transaminases (averaged), BUN – serum blood urea nitrogen, Calcium – serum calcium, Bicarb – serum bicarbonate, Cr – serum creatinine, HCT – hematocrit, INR – international normalized ratio, K+ - serum potassium, PLT – platelet count, PTT – partial thromboplastin time, TBili – total serum bilirubin, WBC – white blood cell count, Stoolþ - post operative bowel movement, BMI – body mass index, HR – heart rate, O2 Sat – pulse oximetry oxygen saturation, RR – respiratory rate, DBP – diastolic blood pressure, SBP – systolic blood pressure, TEMP – body temperature, DxImg – diagnostic imaging in last 24 h, UnivHosp – modifier for geographic site of operation, PreOptimize – dummy variable for hospital stay prior to surgery, LnLOS – natural logarithm of length of stay, LOS – length of stay.

[1] Katz scores are calculated from the Katz Index of Independence in Activities of Daily Living (Gerontological Society of America). 1 point in each of six domains: bathing, dressing, toileting, transferring, continence, and feeding. A total of 6 points indicates complete functional independence, and a total of 0 points indicates no independence.

[2] Variables found to be statistically insignificant with both models are not shown.

[3] Estimated coefficients instead of estimated hazard ratios were reported because of ease of signage interpretation.

3.4 Comparative analysis of discharge and readmission models

The discharge and readmission models were directly compared to identify possible discharge behaviors that were discordant with the goal of reducing surgical readmissions. All of the estimated coefficients used for this analysis – as well as their interpretation – are described in Table 3.2 and Table 3.4.

Table 3.3 and Fig. 3.1 show comparisons of estimated coefficients for standardized variables from the two regression models. Table 3.3 demonstrates the relationship between a clinical variable's normalized discharge regression coefficient and its normalized readmission regression coefficient. Using the patient's reported pain score as an example, the marginal effect on the odds of readmission by a one unit increase in the standardized pain score is estimated to be 17% (the exponential function of 0.158 minus 1) and the estimated marginal effect on the rate of discharge is -6.6% (the exponentional function of -0.068 minus 1). This constellation of coefficients imply that higher pain scores are associated with a greater likelihood of readmission, and that physicians heeded pain concerns by holding discharges on those with higher pain scores.

Table 3.3: **Comparison of the effect of observed clinical variables on predictive models of discharge and readmission.** The sign of the readmission regression coefficient has been reversed for easier direct comparison of the variable's effect on discharging the patient (larger regression coefficient is a higher likelihood of discharge) versus the variable's effect on preventing readmission (larger regression coefficient is a lower likelihood of readmission). Confidence intervals of the coefficients estimates are reported to aid in the interpretation of conventional p-values for large sample sizes.

Qualitative assessmentof discharge decision	Clinical Variable	(A) Discharge Regression			(B) Readmission Regression		
		Est.	95% C.I.		Est.	95% C.I.	
Non-optimal discharge (insufficient attention)	TEMP	0.003	-0.030	0.036	-0.048	-0.005	-0.092 **
	RR	-0.003	-0.031	0.025	-0.110	-0.028	-0.192 ***
	AlkPhos	0.017	-0.023	0.056	-0.075	-0.029	-0.121 ***
	DBP	0.017	-0.008	0.042	-0.065	-0.026	-0.104 ***
Non-optimal discharge (inappropriate effect)	AST/ALT	-0.039	-0.063	-0.016 ***	0.186	0.337	0.035 **
	HighRisk	-0.856	-0.986	-0.726 ***	0.175	0.311	0.039 **
	SBP	-0.024	-0.040	-0.008 ***	0.084	0.133	0.035 ***
	Sodium	-0.061	-0.121	-0.002 **	0.080	0.130	0.031 ***
	Albumin	-0.021	-0.031	-0.012 ***	0.040	0.062	0.018 ***
Optimal Discharge (appropriate effect)	UnivHosp vs Other	-0.469	-0.567	-0.370 ***	-0.372	-0.145	-0.598 ***
	BlackRace	-0.114	-0.164	-0.064 ***	-0.366	-0.133	-0.599 ***
	WhiteRace	-0.046	-0.069	-0.022 ***	-0.279	-0.087	-0.472 ***
	CHF	-0.193	-0.258	-0.128 ***	-0.275	-0.123	-0.427 ***
	DM	-0.038	-0.071	-0.004 **	-0.254	-0.149	-0.358 ***
	IVDU	-0.481	-0.689	-0.273 ***	-0.251	-0.158	-0.343 ***
	Ascites	-0.452	-0.546	-0.358 ***	-0.193	0.019	-0.404 *
	Calcium	0.182	0.108	0.256 ***	0.178	0.236	0.120 ***
	DxImg	-0.650	-1.282	-0.018 **	-0.175	0.020	-0.369 *
	PainScore	-0.068	-0.089	-0.047 ***	-0.158	-0.130	-0.185 ***
	Stool+	0.008	0.005	0.011 ***	0.153	0.292	0.013 **
	HTN	-0.052	-0.086	-0.019 ***	-0.144	-0.027	-0.261 **
	HCT	0.070	0.018	0.121 ***	0.138	0.179	0.097 ***
	CANCER	-0.167	-0.366	0.031 *	-0.130	0.015	-0.276 *
	K+	-0.035	-0.072	0.003 *	-0.130	-0.068	-0.191 ***
	WBC	-0.190	-0.233	-0.147 ***	-0.112	-0.020	-0.203 **
	HR	-0.166	-0.203	-0.130 ***	-0.108	-0.030	-0.187 ***
	BUN	-0.170	-0.228	-0.113 ***	-0.086	0.012	-0.184 *
	TBili	-0.109	-0.151	-0.067 ***	-0.065	0.006	-0.137 *
	Katz Score	0.314	0.271	0.356 ***	0.062	0.098	0.026 ***
	Male Gender	-0.045	-0.096	0.005 *	-0.061	0.004	-0.125 *
	PTT	-0.016	-0.029	-0.004 ***	-0.056	-0.028	-0.084 ***
	PLT	0.075	0.050	0.100 ***	0.042	0.084	0.000 **
Optimal Discharge (unknown effect)	NPO	-2.069	-2.908	-1.229 ***	0.331	0.875	-0.213
	Complexity (RVUs)	-0.198	-0.364	-0.032 **	-0.069	0.027	-0.165
	Reg Admit	-0.098	-0.182	-0.013 **	-0.074	0.204	-0.353
	PtAge	-0.043	-0.072	-0.013 ***	-0.004	0.080	-0.089
	BMI	0.035	0.020	0.051 ***	0.010	0.086	-0.066
	O2Sat	-0.026	-0.039	-0.013 ***	0.026	0.060	-0.007

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

[1] Estimated coefficients instead of estimated hazard ratios were reported because of ease of signage interpretation.

All clinical variables were divided into important qualitative categories based on signage of the estimated coefficients. "Non-optimal discharges" were those in which a variable did not appear to be used effectively in the discharge process. This category of variables was further subdivided into "insufficient attention" (those variables not significantly predictive of discharge but demonstrating an increased risk of readmission) and "inappropriate effect" (those variables that decrease the probability of discharge but were ultimately protective against readmission. The majority of variables fall into the "Optimal Discharge - Appropriate effect" subcategory suggesting that in most cases surgeons were responding to variables predictive of readmission by prolonging their initial hospital stays. There is a final category, "Optimal Discharge – Unknown effect" which highlights the variables that were predictive of discharge but were not found to be significant predictors of readmission. Based on this study's methodology, it is impossible to interpret whether these variables were being appropriately selected for by surgeons or if these clinical delays were unnecessary for reducing readmissions.

For variables statistically significant in both models, Fig. 3.1 graphically demonstrates the degree of alignment between factors that are predictive of discharge and those that are predictive of readmission. Quadrants I and III of the figure show concordance between current discharge behaviors and reducing surgical admissions. For example, a patient carrying a diagnosis of congestive heart failure (CHF) is both less likely to be discharged and more likely to be readmitted. In contrast, Quadrants II and IV represent discordant results between the models. Using liver function as an example, elevated liver transaminases lead to less likely discharge even though that patient is less likely to be readmitted. On the other hand, patients with an isolated elevated alkaline phosphatase are more likely to be discharged and more likely to be readmitted. In addition, there are several effect size discrepancies noted. Patient race appears to have an important impact on readmission with white patients substantially more likely to be readmitted than non-white patients, but discharge decision-making does not appear to be significantly affected by it. Conversely, a history of intravenous drug use (IVDU) decreases likelihood of discharge but this effect may be disproportionate to the effect of IVDU on readmission risk.

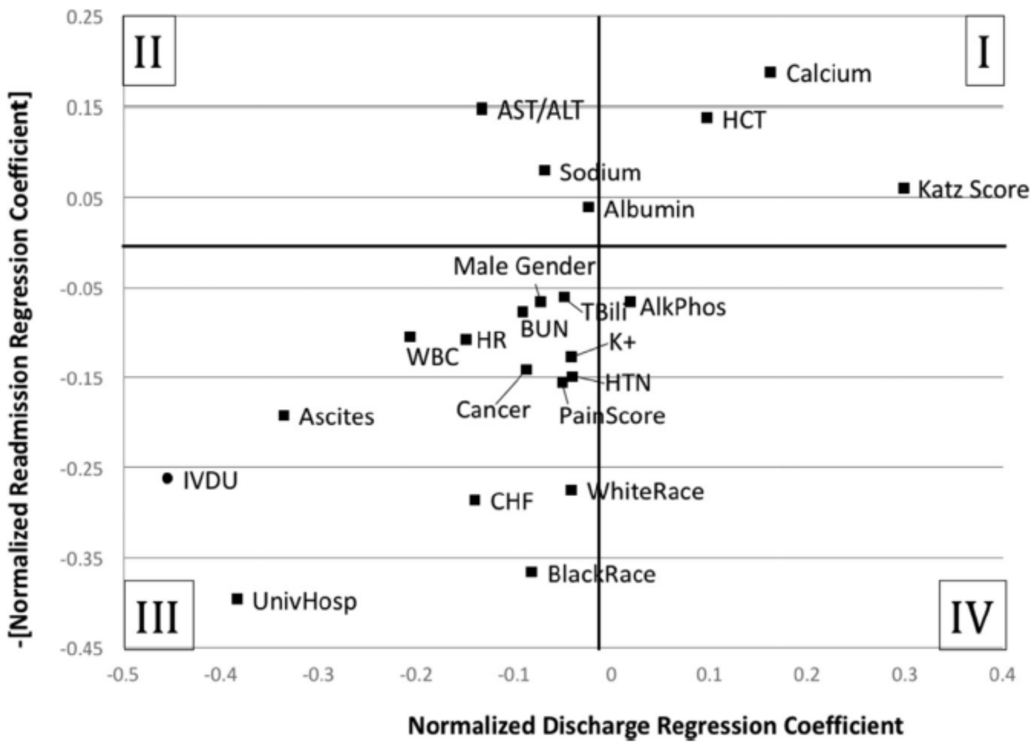


Figure 3.1: **Comparison of the effect of observed clinical variables on predictive models of discharge and readmission.** The x-axis plots the normalized regression coefficient of a time-to-event model of the likelihood of discharge for a given day's observed clinical variables. The yF-axis plots the normalized coefficient of a logit regression model of the likelihood of 30-day readmission. The sign of the readmission regression coefficient has been reversed for easier direct comparison of the variable's effect on discharging the patient (larger regression coefficient is a higher likelihood of discharge) versus the variable's effect on preventing readmission (larger regression coefficient is a lower likelihood of readmission). In its current projection, Quadrant I demonstrates variables that both increase the likelihood of discharge and reduce the risk of readmission; Quadrant III demonstrates variables that decrease the likelihood of discharge and increase the risk of readmission. Quadrants II and IV represent discordance between behavior and readmission with the former indicating variables that increase length of stay but reduce the risk of readmission and the latter indicating variables that decrease length of stay but increase risk of readmission. Variables found to be statistically insignificant with either model are not shown.

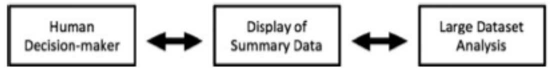


Figure 3.2: **Human-machine interface of decision-support tools.** This schematic illustrates that the ways in which patient data are presented to clinical decision-makers may be as important to the successful implementation of a decision-support tool as the underlying analytical methodology.

3.5 Discussion

This study illustrates the feasibility of using large patient data sets to construct predictive models for readmission that highlight clinical data that is not conventionally used by surgeons when making the decision to discharge a patient from the hospital after an inpatient surgical procedure.

Tables 3.2 and 3.3 with Fig. 3.1 highlight the variables found to be significant in a logit regression model for readmission, demonstrating that variation in a number of variables may be drivers of – or at least likely indicators of a higher risk of readmission. This predictive model of readmissions was also validated with standard statistical techniques demonstrating good discrimination (C statistics of 0.82 and 0.80) that is better than many discrimination measures reported in prior readmission modeling studies. (Kansagara et al. (2011)) When this predictive model was compared with the clinical predictors of discharge, the observed differences in the two models highlighted opportunities for improving discharge decision-making.

Currently observed discharge behaviors might not be fully aligned with current readmission reduction goals. Table 3.3 highlights variation in predictors of discharge versus their effect on readmission. What surgeons practice does not appear to always align with reducing readmissions (Table 3.2, Table 3.4, Table 3.3, and Fig. 3.1). For example, variables that were not significantly associated with surgeons' discharge decisions but positively associated with the likelihood of being readmitted suggest that surgeon behavior may be missing important readmission-reducing data trends. It is possible that some of these less carefully considered variables (e.g., diastolic blood pressure, respiratory rate, liver function tests) are picking up subtle variations - so-called "microtrends" - in clinical trajectories that are not traditionally considered to be relevant to overall discharge decision-making. Perhaps readmissions could be reduced by using a data-driven predictive model that incorporates these less commonly considered clinical data points. This discordance also highlights the potential for poor acceptance of discharge decision-support tools that utilize statistical predictors of readmission that diverge from factors that surgeons believe to be important predictors of a successful discharge. For example, a decision-support tool may highlight a patient's race as an indication that the patient may require more intensive discharge planning with the potential for prolonged index admission. If the surgeon does not intrinsically believe that patient characteristic is relevant for discharge decision-making, he or she may discount the tool's recommendations. Creative strategies to incorporate the best data analysis with display modes that overcome reluctance to use less intuitive but data-driven predictive models are a necessary step for the promotion of such decision-support technologies (Fig. 3.2).

Previous efforts to reduce readmission have started by identifying processes and bundling strategies in the pre- and post-operative period. (Kassin et al. (2012); Halverson et al. (2014)) An alternative but not yet proven strategy suggested for reducing surgical readmissions is the risk stratification of postoperative patients and enhanced discharge decision-making. (Kohlnhofer et al. (2014); Leeds et al. (2013))Each patient is a unique individual, and the patient's physician must retain responsibility for the decision to discharge the patient from the hospital. But these central features of the physician-patient relationship are not at direct odds with the need for further decision-support in the discharge process.

An innovation with promise for promoting evidence-based discharge decision-making involves the development of decision-support software. Expanding on the predictive readmission modeling discussed above, the authors believe that alternative, support-vector regression models that can predict the probability of patient readmission continuously over a patient's hospital stay would be useful to optimize clinicians' discharge decisions. Ultimately, this more comprehensive model could provide dynamically updated marginal probabilities of patient readmission within 30 days of discharge for a specific patient using real-time clinical, demographic, and census data.

Fig. 3.3 provides an example of a comprehensive predictive model illustrating how a patient's LOS may have been unnecessarily prolonged. Although this patient stayed in the hospital for 17 days, the model predicts that the likelihood of readmission is not statistically different from day 12 onward. This example illustrates how a discharge decision might have been improved by the use of a decision-support algorithm.

The algorithm also provides the foundation for a decision-support tool by treating the patient's measured characteristics as one "patient observation" within a large sample of archived patients with similar measured characteristics but known outcomes. In this way, a current discharge decision can be informed by the aggregated experience with thousands of similar patients with known histories. Such a decision-support tool provides a statistically informed answer to the central question: "If this patient is discharged today, what is the likelihood of readmission within 30 days?"

3.5.1 Limitations

The findings reported here are not without interpretative limitations. First, this study was performed with surgeons and patient data from a single institution limiting the generalizability of the variables identified. Importantly, this institution was a relative under-performer compared to its peers with a length of stay index of 1.12 versus a teaching hospital baseline of 1.06 and a readmission rate of 5.26% versus 5.24% (unpublished data, University Health-System Data, accessed 2015). Therefore, this institution may have specific processes and patient factors that lead it to benefit more than average from clinical decision support tells that help optimize patient care. It is likely that institutions already out-performing their peers on quality metrics may demonstrate a different set of predictors than under-performers. It is also possible that outperformers may benefit less from such decision support optimization with limited ability to incrementally improve quality performance further. Moreover, the group of procedures included in this analysis was predominantly gastrointestinal surgery cases limiting generalizability to other surgical procedures. However, we argue that the analytical exercise described in this study is not meant to be directly portable to other clinical sites but instead demonstrates an approach to optimizing decision support tools for the specific patients and surgeon preferences of each health care environment. It is very likely that the predictors of readmission will vary with different populations of surgeons and patients. We believe site-specific customization of decision-support tools will be increasingly important with increasing use of these technologies, and methodologies – rather than ready-made decision tools – will be the currency of future large-scale quality improvement interventions.

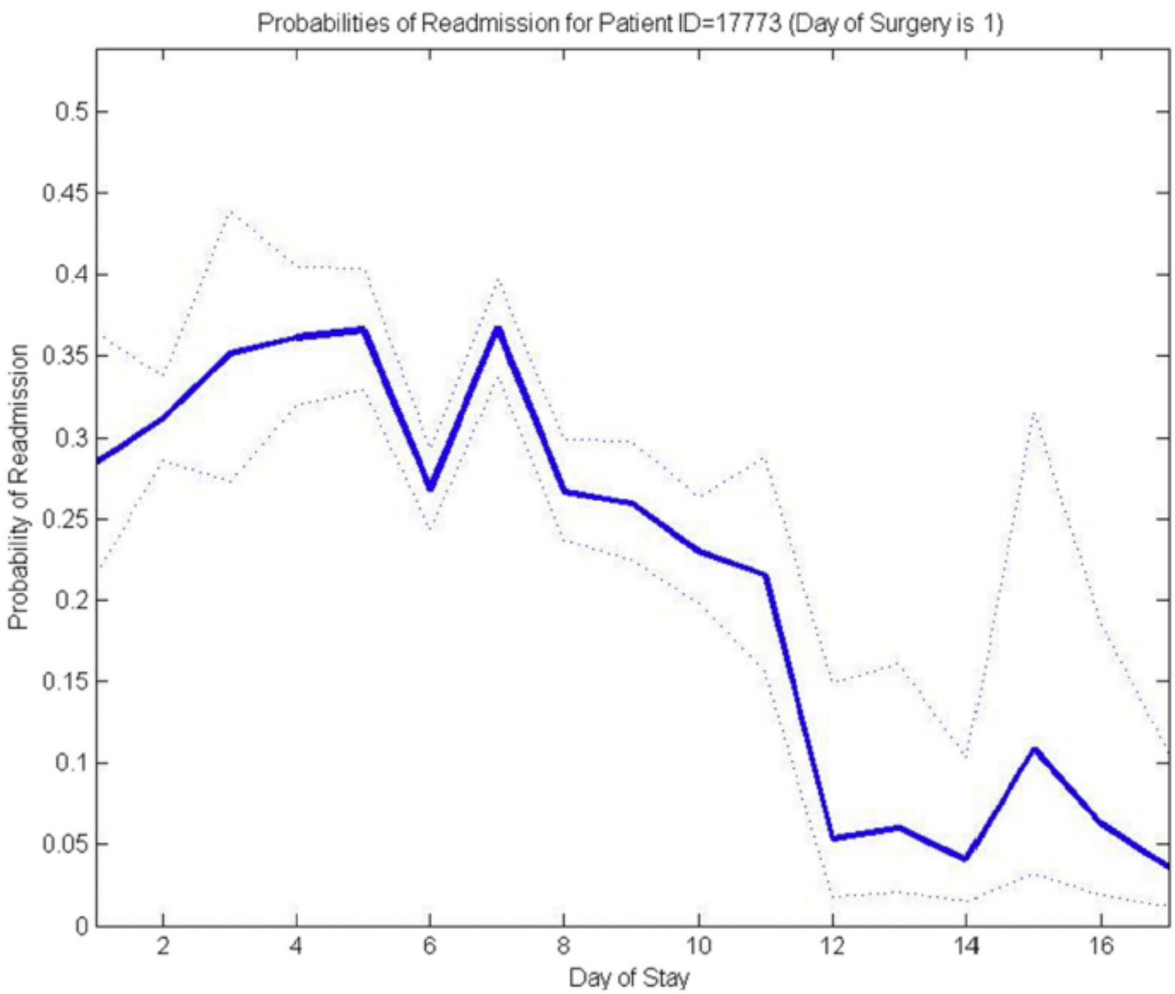


Figure 3.3: **A regression-based algorithm predicting daily risk of readmission for a sample inpatient.** The x-axis indicates elapsed days since surgery (i.e., length of stay) and the y-axis shows predicted risk of readmission if discharged on that day. The solid line represents the point estimates of readmission probabilities generated by the regression model. The dotted lines represent the 80% confidence intervals. For this specific graph, estimates were run from virtual Day 2 of admission because of the extreme unlikelihood of next day discharge for a complex surgical patient.

Another limitation of this study is the imperfect comparison between what was observed at discharge and these variables' subsequent impact on readmission. A variable that the regression model found to be relatively unimportant is not necessarily unimportant for discharge but simply may indicate observations of a variable that fall within such a narrow range for all discharges as to appear insignificant.

Arguing that all of the discordance between surgeon's preferences and the predictive model is the result of human decision-making errors would be an over-interpretation of this analysis. Instead, what the authors choose to highlight is that the discordance between surgeon behaviors and readmission-optimized behaviors based on a predictive model highlight variables for closer examination to determine marginal improvements in discharge decision-making. For example, the discharge regression model suggests that a number of clinical variables have an over-emphasized effect on discharge when compared to their effect on readmission (Table 3.3). Perhaps the proper decision-support tool could better align these effects by reducing selected variables' impact on discharge decision-making.

3.6 Conclusions

Predictive models suggest considerable discordance exists between how surgeons behave in practice and what may optimize discharge decision-making and limit subsequent hospital readmission. The real-world application of such a predictive model is a question in itself. Behavioral experiments are necessary to ascertain how and whether medical decision makers will use decision support software. If so, is the decision support tool effective in reducing readmissions? These further issues will need to be addressed through ongoing research.

Table 3.4: **Normalized effect of 23 procedure group variables on discharge and readmission (n = 20,970 patients with 115,976 daily observed clinical measures; 1565 readmitted patients with 20,560 daily observed clinical measures).** Column A reports the normalized regression coefficient for each procedure for a time-to-event model analysis predicting daily likelihood of discharge based on actual discharges of the patient population. A positive value indicates the presence or increased magnitude of the variable increases the likelihood of discharge, and vice versa. Column C reports the normalized coefficient for each clinical variable for a logit regression model predicting likelihood of 30-day readmission. A positive value indicates the presence or increased magnitude of the variables increases the likelihood of readmission. Omitted coefficients are those that were not statistically supported by one of the two models.

Procedure coefficient	Readmission regression coefficient	%95 C.I.		p
Adrenalectomy	-0.78	-1.071	-0.488	0.000 ***
Appendectomy	-1.048	-1.452	-0.644	0.000 ***
Hepato-Jejuno/Duodeno	-0.746	-1.101	-0.391	0.000 ***
Pancreatic Debride	-1.077	-1.413	-0.741	0.000 ***
Esophagectomy	-0.948	-1.289	-0.608	0.000 ***
Gastrectomy	-0.652	-0.946	-0.359	0.000 ***
Gastro-Jejuno	-0.785	-1.114	-0.455	0.000 ***
Gastrorraphy	-1.105	-1.447	-0.763	0.000 ***
Parathyroidectomy	-0.94	-1.300	-0.579	0.000 ***
Thyroidectomy	-1.454	-1.888	-1.019	0.000 ***
Bariatric Surgery	-0.656	-1.052	-0.259	0.001 ***
Cholecystectomy	-0.628	-0.987	-0.269	0.001 ***
GI Tumor Excision	-0.567	-0.908	-0.227	0.001 ***
Lysis of Adhesions	-0.79	-1.247	-0.333	0.001 ***
Splenectomy	0.515	0.203	0.827	0.001 ***
Colectomy	-0.399	-0.657	-0.141	0.002 ***
Ing-Fem Herniorraphy	-0.771	-1.258	-0.284	0.002 ***
Hepatectomy	-0.668	-1.099	-0.237	0.002 ***
PEH Repair	-0.9	-1.583	-0.218	0.010 **
Ventral Herniorraphy	-0.323	-0.612	-0.033	0.029 **
Mastectomy	-0.542	-1.072	-0.011	0.045 **
Small Bowel Resect	-0.342	-0.676	-0.007	0.045 **
Soft Tissue Debride	0.323	0.004	0.642	0.047 **
GI Abscess Drain	0.233	-0.032	0.498	0.085 *

***p<0.01, **p<0.05, *p<0.1.

Financial support

This work was supported in part by National Institutes of Health/National Institute of Aging Grant 1RC4AG039071 (to J.F.S. and J.C.C).

Potential conflicts of interest

V.S., J.C.C., K.E.S., and J.F.S. report owning equity interests in 4C Health Analytics, Inc., a start-up company that may in future market healthcare IT products.

Acknowledgments

The authors appreciate the critical review of a draft version of this manuscript performed by Dr. Sandra Zaeh and econometric methodological assistance by Dr. Shiferaw Gurmu.

APPENDIX A

APPENDIX FIGURES

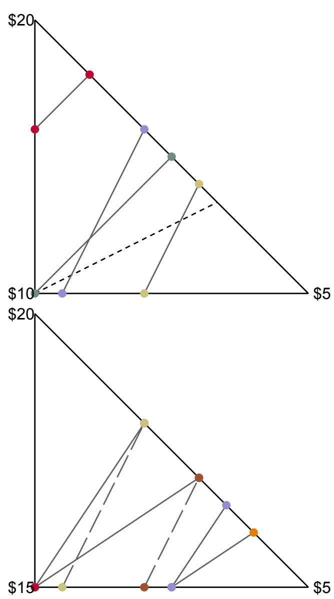


Figure A.1: 20 Lottery Pairs Shown in MM Triangles

Figure A.2: Difference in Signaling Frequencies per Seller

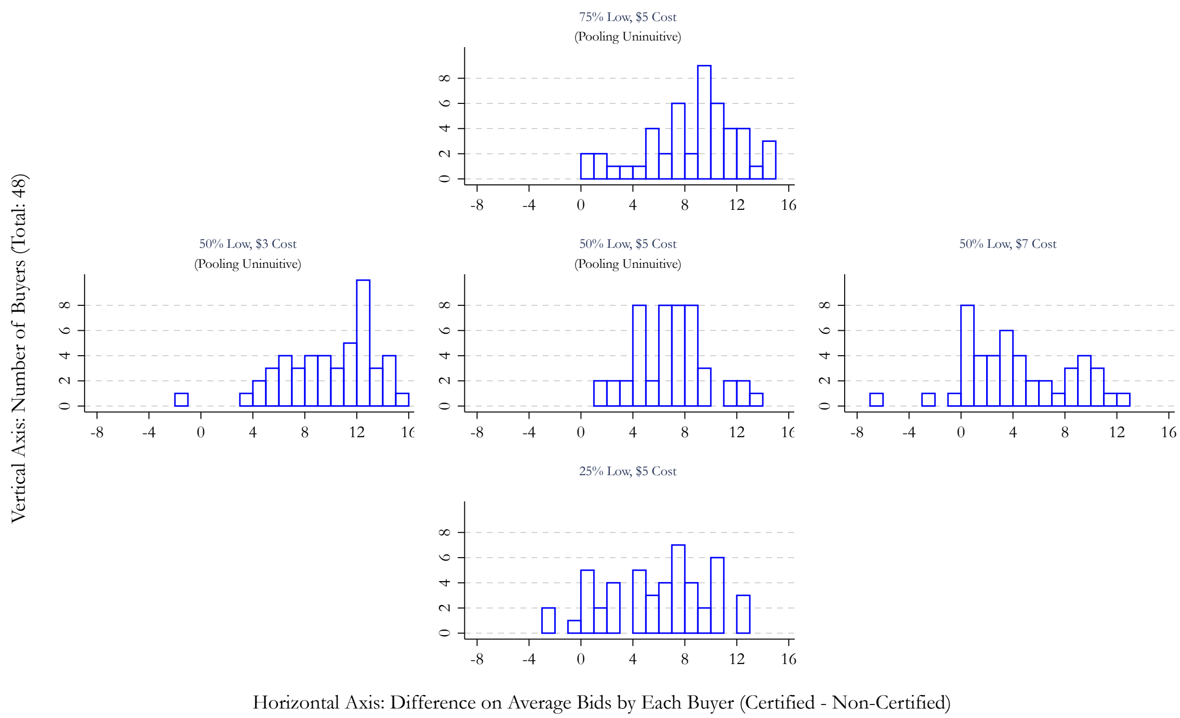


Figure A.3: Difference in Average Bids per Buyer

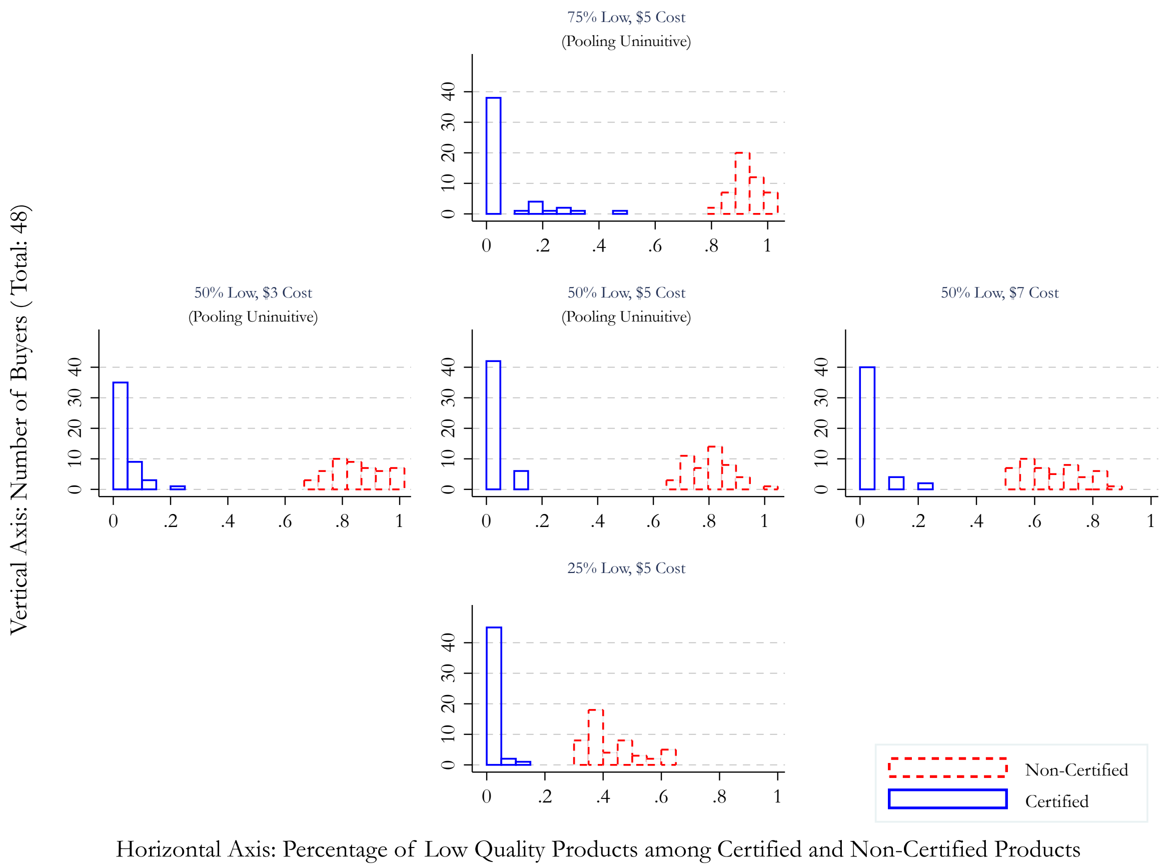


Figure A.4: Percentage of Low Quality Products Encountered per Buyer (By Product Certification)

APPENDIX B

APPENDIX TABLES

Table B.1: Demographics of Subjects in Each Treatment

Treatment		1	2	3	4	5
Total Subjects		48	48	48	48	48
Female		26	20	32	30	31
No Experience		23	19	25	27	16
Age		19.7	20.6	20.4	19.7	19.9
Race	African American	30	26	30	31	26
	White	5	3	5	6	6
	Asian	11	14	7	5	8
	Hispanic	1	4	2	1	2
	Other or Prefer Not to Answer	1	1	4	5	6
GPA	< 2.0	1	0	2	0	0
	2.0-2.99	7	10	10	17	11
	3.0-3.49	21	18	21	18	19
	3.5-4.3	19	20	15	13	18

Table B.2: Comparison of Signaling Rates in the First and Last Blocks of Each Treatment

Treatment	Pooling is Unintuitive	Signaling Rates (%)								
		High Type S_H			Low Type S_L			Difference S_{DIFF}		
		1st Block	4th Block	Difference	1st Block	4th Block	Difference	1st Block	4th Block	Difference
1	Yes	75.0	90.6	15.6	13.5	6.3	-7.3	61.5	84.4	22.9 **
2	Yes	70.3	85.9	15.6 ***	12.5	9.4	-3.1	57.8	76.6	18.8 **
3	No	66.7	67.7	1.0	15.6	9.4	-6.3	51.0	58.3	7.3
4	Yes	76.6	92.2	15.6 **	26.6	23.4	-3.1	50.0	68.8	18.8 ***
5	No	53.1	43.8	-9.4	12.5	4.7	-7.8 **	40.6	39.1	-1.6

Table B.3: Comparison of Average Bids in the First and Last Blocks of Each Treatment

Treatment	Pooling is Unintuitive	Average Bids ($)								
		Non-Certified B_{NC}			Certified B_C			Difference B_{DIFF}		
		1st Block	4th Block	Difference	1st Block	4th Block	Difference	1st Block	4th Block	Difference
1	Yes	7.74	8.97	1.24 ***	16.05	18.92	2.87 ***	8.30	9.94	1.65
2	Yes	9.02	9.49	0.47 ***	15.06	17.66	2.60 **	6.03	8.17	2.13 ***
3	No	10.61	12.22	1.61 ***	15.98	18.28	2.30 ***	5.36	6.05	0.69
4	Yes	9.16	9.18	0.02	17.88	20.38	2.50 ***	8.73	11.20	2.47 ***
5	No	8.51	10.19	1.68 ***	13.25	15.33	2.08 ***	4.76	5.14	0.38

Table B.4: Regression Analysis of Signaling Behavior

1=Signal	Low Quality Products				High Quality Products			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treament 1 (75%, $5)	-0.254	-0.255	-0.188	-0.155	0.285	0.286	0.281	0.231
	(0.315)	(0.318)	(0.318)	(0.337)	(0.361)	(0.361)	(0.362)	(0.365)
Treament 3 (25%, $5)	0.0277	0.0278	0.0775	0.0702	-0.887***	-0.889***	-0.892***	-0.928***
	(0.326)	(0.327)	(0.331)	(0.350)	(0.306)	(0.306)	(0.308)	(0.314)
Treament 4 (50%, $3)	0.233	0.234	0.346	0.377	0.603*	0.604*	0.597*	0.741**
	(0.290)	(0.291)	(0.303)	(0.319)	(0.321)	(0.322)	(0.325)	(0.346)
Treament 5 (50%, $7)	-0.715**	-0.718**	-0.725**	-0.689*	-1.374***	-1.377***	-1.378***	-1.329***
	(0.352)	(0.352)	(0.339)	(0.370)	(0.316)	(0.317)	(0.317)	(0.327)
Period		-0.0206***	-0.0211***	-0.0212***		0.0103	0.0103	0.00924
		(0.00733)	(0.00717)	(0.00705)		(0.00745)	(0.00745)	(0.00765)
Number of Safe Choices			-0.0850**	-0.0752**			0.00558	0.0129
			(0.0344)	(0.0333)			(0.0278)	(0.0294)
Age				-0.154**				-0.0207
				(0.0741)				(0.0460)
Male				0.418*				0.364
				(0.219)				(0.232)
Experienced Subjects				-0.242				0.0445
				(0.210)				(0.214)
Sophomore				1.981**				0.640
				(0.885)				(0.835)
Junior				0.354				0.308
				(0.314)				(0.325)
Senior				0.725				0.404
				(0.448)				(0.379)
Graduate Students				0.206				0.128
				(0.312)				(0.277)
GPA				0.0556				0.362***
				(0.151)				(0.131)
African American				0.0994				0.165
				(0.299)				(0.295)
Hispanic				0.599				0.151
				(0.413)				(0.436)
Asian				0.989***				0.317
				(0.379)				(0.434)
Other				-0.135				0.247
				(0.388)				(0.459)
Smoke				0.185				0.212
				(0.262)				(0.353)
Youngest Child				0.289				0.0461
				(0.304)				(0.272)
Middle Child				0.541				0.879
				(0.477)				(0.614)
Oldest Chid				0.202				0.343
				(0.274)				(0.282)
Constant	-1.600***	-1.272***	-0.393	1.761	1.343***	1.175***	1.116***	-0.256
	(0.209)	(0.235)	(0.410)	(1.552)	(0.239)	(0.272)	(0.403)	(1.083)
Observations	1,280	1,280	1,280	1,280	1,280	1,280	1,280	1,280

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.5: Regression Analysis of Bidding Behavior

Y: Bids	Non-Certified Products				Certified Products			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treament 1 (75%, $5)	-0.799***	-0.806***	-0.813***	-0.914***	1.298**	1.278**	1.212*	1.189*
	(0.277)	(0.276)	(0.272)	(0.283)	(0.641)	(0.638)	(0.645)	(0.635)
Treament 3 (25%, $5)	2.912***	2.908***	2.908***	2.790***	1.886***	1.910***	1.889***	1.940***
	(0.482)	(0.480)	(0.479)	(0.484)	(0.642)	(0.640)	(0.633)	(0.569)
Treament 4 (50%, $3)	-0.0157	-0.0184	-0.0240	-0.244	3.158***	3.180***	3.131***	3.264***
	(0.289)	(0.290)	(0.287)	(0.304)	(0.651)	(0.655)	(0.659)	(0.607)
Treament 5 (50%, $7)	0.113	0.0800	0.0788	0.0359	-1.121	-0.920	-0.898	-0.530
	(0.334)	(0.331)	(0.331)	(0.340)	(0.775)	(0.769)	(0.773)	(0.765)
Period		0.0479***	0.0479***	0.0465***		0.112***	0.112***	0.115***
		(0.00796)	(0.00796)	(0.00786)		(0.0157)	(0.0157)	(0.0159)
Number of Safe Choices			0.00764	-0.00878			0.0938	0.0914
			(0.0282)	(0.0295)			(0.0573)	(0.0569)
Age				0.0102				0.248***
				(0.0428)				(0.0777)
Male				0.0652				1.068**
				(0.264)				(0.431)
Experienced Subjects				-0.385				-0.295
				(0.251)				(0.407)
Sophomore				-0.407				-2.840*
				(0.727)				(1.503)
Junior				0.00606				-0.249
				(0.365)				(0.596)
Senior				-0.444				-1.138
				(0.420)				(0.857)
Graduate Students				0.217				-1.889***
				(0.265)				(0.603)
GPA				-0.0635				0.202
				(0.157)				(0.294)
African American				0.720**				0.302
				(0.283)				(0.529)
Hispanic				0.256				-0.244
				(0.634)				(0.954)
Asian				0.235				0.342
				(0.417)				(0.910)
Other				0.299				0.325
				(0.302)				(0.839)
Smoke				0.349				0.0819
				(0.355)				(0.605)
Youngest Child				-0.256				0.304
				(0.299)				(0.575)
Middle Child				-0.00454				1.590
				(0.437)				(0.986)
Oldest Chid				0.00619				0.0450
				(0.288)				(0.523)
Constant	9.146***	8.372***	8.288***	8.316***	16.07***	14.15***	13.11***	7.813***
	(0.250)	(0.299)	(0.462)	(1.052)	(0.431)	(0.551)	(0.818)	(2.077)
Observations	3,456	3,456	3,456	3,456	1,664	1,664	1,664	1,664
R-squared	0.172	0.194	0.194	0.214	0.090	0.136	0.141	0.191

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

APPENDIX C

APPENDIX PROOFS

In this appendix, I will discuss the three potential pure-strategy equilibria in the model outlined in Section 1.2, and derive the conditions in the propositions and corollaries. First, given the wage schedule based on employers' beliefs, define the wage difference as

$$w(1) - w(0) = (1-\mu_1)\theta_h + \mu_1\theta_l - (1-\mu_0)\theta_h - \mu_0\theta_l = (\mu_0 - \mu_1)(\theta_h - \theta_l).$$

Pooling not to pursue $e_h = e_l = 0$

In this case, no workers choose to pursue the degree, and all workers will be on the job market without a degree. As a result, on-equilibrium belief μ_0 will be the same as prior distribution $\mu_0 = \mu$; employers will not observe any workers with the degree in the market and cannot use Bayes' rule to update μ_1, so out-of-equilibrium belief μ_1 can take any value in $[0,1]$. The wage schedule will be

$$\begin{aligned} w(0) &= (1-\mu)\ \theta_h + \mu\ \theta_l \\ w(1) &= (1-\mu_1)\ \theta_h + \mu_1\theta_l\ \in [\theta_l, \theta_h]. \end{aligned}$$

Workers' decisions to not pursue the degree will be supported by this wage schedule if

$$(1-\lambda_h)w(1) + \lambda_h w(0) - c_h \leq w(0) \qquad \text{High Type}$$

$$(1-\lambda_l)w(1) + \lambda_l w(0) - c_l \leq w(0) \qquad \text{Low Type}$$

The two incentive constraints can be rearranged as:

$$(1-\lambda_h)(w(1)-w(0)) \leq c_h$$
$$(1-\lambda_l)(w(1)-w(0)) \leq c_l,$$

here $w(1)-w(0)=(\mu-\mu_1)(\theta_h-\theta_l)$. If $w(1)\leq w(0)$ – i.e., $1\geq\mu_1\geq\mu$ – the L.H.S of both constraints is non-positive, and given that c_h and c_l are strictly positive, the constraints are satisfied; if $w(1)>w(0)$ – i.e., $\mu>\mu_1\geq 0$ –low type's constraint will be implied by the high type's. Solving the high type's constraint, we have:

$$\begin{aligned}
w(1)-w(0) &\leq \frac{c_h}{1-\lambda_h} \\
(\mu-\mu_1)(\theta_h-\theta_l) &\leq \frac{c_h}{1-\lambda_h} \\
\mu-\mu_1 &\leq \frac{c_h}{(1-\lambda_h)(\theta_h-\theta_l)} \\
\mu_1 &\geq \mu-\frac{c_h}{(1-\lambda_h)(\theta_h-\theta_l)} \\
\mu_1 &\geq \mu-\frac{1}{\theta_h-\theta_l}\cdot\frac{c_h}{1-\lambda_h}
\end{aligned}$$

Together with the pre-condition $\mu>\mu_1\geq 0$, we have

$$\mu>\mu_1\geq\mu-\frac{1}{\theta_h-\theta_l}\cdot\frac{c_h}{1-\lambda_h}$$

Combining with the case $1\geq\mu_1>\mu$, the range of supporting out-of-equilibrium beliefs is

$$1>\mu_1\geq\mu-\frac{1}{\theta_h-\theta_l}\cdot\frac{c_h}{1-\lambda_h}$$

Consider the cases in which the low type's constraint $(1-\lambda_l)(w(1)-w(0)) \leq c_l$ is always satisfied no matter what μ_1 is. That is, for any $\mu_1 \in [0,1]$ we should have

$$\begin{aligned}
w(1)-w(0) &\leq \frac{c_l}{1-\lambda_l} \\
(\mu-\mu_1)(\theta_h-\theta_l) &\leq \frac{c_l}{1-\lambda_l} \\
\mu-\mu_1 &\leq \frac{c_l}{(1-\lambda_l)(\theta_h-\theta_l)} \\
\mu_1 &\geq \mu - \frac{c_l}{(1-\lambda_l)(\theta_h-\theta_l)}, \quad \forall \mu_1 \in [0,1] \\
0 &\geq \mu - \frac{c_l}{(1-\lambda_l)(\theta_h-\theta_l)} \\
\frac{c_l}{\theta_h-\theta_l} &\geq \mu(1-\lambda_l)
\end{aligned}$$

Therefore, if $\frac{c_l}{\theta_h-\theta_l} \geq \mu(1-\lambda_l)$, deviating to $e_l = 1$ is never profitable for the low type no matter how favorable the employer's out-of-equilibrium belief is. At the same time, if $\frac{c_h}{\theta_h-\theta_l} < \mu(1-\lambda_h)$, then deviating to $e_h = 1$ is profitable for the high-type worker under some favorable belief of employer $[0, \mu - \frac{c}{(1-\lambda_h)(\theta_h-\theta_l)}]$. So, by the Cho-Kreps intuitive criterion, the employer's out-of-equilibrium belief should be updated to $\mu_1 = 0$, and he will offer a wage of θ_h, which will attract the high-type worker away and disturb this pooling equilibrium.

Separating $e_h = 1$ and $e_l = 0$

In this case, the high-type workers will pursue the degree while the low type will not. As a result, employers are able to verify productivities for workers with and without the degree and to update their beliefs using Bayes rule:

$$\mu_0 = \frac{\mu}{\mu+\lambda_h(1-\mu)}, \quad \mu_1 = 0$$

Given the updated belief, the wage schedule will be

$$w(0) = \frac{\lambda_h(1-\mu)}{\mu+\lambda_h(1-\mu)} \cdot \theta_h + \frac{\mu}{\mu+\lambda_h(1-\mu)} \cdot \theta_l, \quad w(1) = \theta_h$$

And for the market to be in equilibrium, the wage schedule will need to sustain workers' choices:

$$(1-\lambda_h)w(1) + \lambda_h w(0) - c_h \geq w(0) \qquad \text{High Type}$$

$$(1-\lambda_l)w(1) + \lambda_l w(0) - c_l \leq w(0) \qquad \text{Low Type}$$

Rearrange the incentive constraints as:

$$(1-\lambda_h)(w(1) - w(0)) \geq c_h \qquad \text{High Type}$$

$$(1-\lambda_l)(w(1) - w(0)) \leq c_l, \qquad \text{Low Type}$$

Where $w(1) - w(0) = \mu_0(\theta_h - \theta_l) = \frac{\mu}{\mu+\lambda_h(1-\mu)}(\theta_h - \theta_l)$. Further simplifying the constraints, we have:

$$\frac{c_h}{1-\lambda_h} \leq w(1) - w(0) \leq \frac{c_l}{1-\lambda_l}$$
$$\frac{c_h}{1-\lambda_h} \leq \frac{\mu}{\mu+\lambda_h(1-\mu)}(\theta_h - \theta_l) \leq \frac{c_l}{1-\lambda_l}$$

So separating equilibrium exists if and only if the parameters satisfy

$$\frac{c_h}{1-\lambda_h} \leq \frac{\mu}{\mu+\lambda_h(1-\mu)}(\theta_h - \theta_l) \leq \frac{c_l}{1-\lambda_l}$$

Pooling to pursue $e_h = e_l = 1$

Finally, consider the case in which both types choose to pursue the degree. Although both types of workers choose to pursue the degree, due to a positive failure rate, the employer will still observe $e_o = 0$ and will be able to update his belief μ_0. In this situation, the employer's beliefs given observing $e_0 = 1$ and $e_o = 0$ are

$$\mu_1 = \frac{(1-\lambda_l)\mu}{(1-\lambda_l)\mu + (1-\lambda_h)(1-\mu)}, \quad \mu_0 = \frac{\lambda_l \mu}{\lambda_l \mu + \lambda_h (1-\mu)}$$

Given the beliefs, the wage schedule will be

$$w(e) = \begin{cases} (1-\mu_0)\theta_h + \mu_0 \theta_l, & e = 0 \\ (1-\mu_1)\theta_h + \mu_1 \theta_l, & e = 1 \end{cases}$$

Note that since we have $\lambda_h < \lambda_l$, it is always the case that $\mu_0 > \mu_1$, as shown below, so $w(1) - w(0) = (\mu_0 - \mu_1)(\theta_h - \theta_l) > 0$.

$$\begin{aligned} \mu_0 - \mu_1 &= \frac{\lambda_l \mu}{\lambda_l \mu + \lambda_h (1-\mu)} - \frac{(1-\lambda_l)\mu}{(1-\lambda_l)\mu + (1-\lambda_h)(1-\mu)} \\ &= \frac{\lambda_l \mu[(1-\lambda_l)\mu + (1-\lambda_h)(1-\mu)] - (1-\lambda_l)\mu[\lambda_l \mu + \lambda_h(1-\mu)]}{[(1-\lambda_l)\mu + (1-\lambda_h))(1-\mu)][(\lambda_l \mu + \lambda_h(1-\mu))]} \\ &= \frac{\lambda_l \mu[(1-\lambda_h)(1-\mu)] - (1-\lambda_l)\mu[\lambda_h(1-\mu)]}{[(1-\lambda_l)\mu + (1-\lambda_h)(1-\mu)][(\lambda_l \mu + \lambda_h(1-\mu))]} \\ &= \frac{\mu(1-\mu)[\lambda_l(1-\lambda_h) - (1-\lambda_l)\lambda_h]}{[(1-\lambda_l)\mu + (1-\lambda_h)(1-\mu)][(\lambda_l \mu + \lambda_h(1-\mu))]} \\ &= \frac{\mu(1-\mu)(\lambda_l - \lambda_h)}{[(1-\lambda_l)\mu + (1-\lambda_h)(1-\mu)][\lambda_l \mu + \lambda_h(1-\mu)]} > 0 \end{aligned}$$

Workers' incentive constraints can be rearranged as:

$$(1-\lambda_h)(w(1) - w(0)) \geq c_h \qquad \text{(For High Productive Worker)}$$

$$(1-\lambda_l)(w(1) - w(0)) \geq c_l, \qquad \text{(For Low Productive Worker)}$$

where $w(1) - w(0) = (\mu_0 - \mu_1)(\theta_h - \theta_l) > 0$. Since we also have $\lambda_h < \lambda_l$ and $c_h < c_l$, the high type's constraint is implied by the low type's:

$$
\begin{aligned}
(1-\lambda_l)(w(1) - w(0)) &\geq c_l \\
(1-\lambda_l)(\mu_0 - \mu_1)(\theta_h - \theta_l) &\geq c_l \\
(\mu_0 - \mu_1)(\theta_h - \theta_l) &\geq \frac{c_l}{1-\lambda_l}
\end{aligned}
$$

When this inequality is not satisfied, pooling at $e = 1$ can be avoided; that is:

$$
\begin{aligned}
&\frac{c_l}{1-\lambda_l} \geq (\mu_0 - \mu_1)(\theta_h - \theta_l) \\
&\text{where } \mu_1 = \frac{(1-\lambda_l)\mu}{(1-\lambda_l)\mu + (1-\lambda_h)(1-\mu)}, \quad \mu_0 = \frac{\lambda_l \mu}{\lambda_l \mu + \lambda_h (1-\mu)}
\end{aligned}
$$

Comparing this condition with the condition on $\frac{c_l}{1-\lambda_l}$ for existence of a separating equilibrium, we have

$$
\mu_0 - \mu_1 \leq \mu_0 = \frac{\mu}{\mu + \frac{\lambda_h}{\lambda_l}(1-\mu)} \leq \frac{\mu}{\mu + \lambda_h(1-\mu)}
$$

Since $\lambda_l < 1$, we have $\frac{\lambda_h}{\lambda_l} > \lambda_h$. This means that, when the separating strategy $e_h = 1, e_l = 0$ is an equilibrium strategy, pooling to pursue cannot be an equilibrium.

The last case is the most wasteful; to show this, denote the beliefs in this pooling equilibrium μ_0^{Pool} and μ_1^{Pool}. Educational attainment, in this pooling equilibrium, can act as a weak signal of a worker's type since we have $\mu_0^{Pool} > \mu > \mu_1^{Pool}$; that is, the degree status is negatively correlated with a worker's likelihood to be the low type. However, recall that in the separating equilibrium, the negative relationship is stronger, and, therefore, a degree is a stronger signal since

$$
\mu_0^{Sep} = \frac{\mu}{\mu + \lambda_h(1-\mu)} > \mu_0^{Pool}, \quad \mu_1^{Sep} = 0 < \mu_1^{Pool}
$$

For given c_h and c_l, we conclude that pooling to pursue is less efficient than the separating equilibrium, since everyone incurs the cost, but the degree does not convey as much information.[1]. The inefficient pooling equilibrium can be *avoided* by raising the low type's dropout risk.

Sequential Equilibrium

Perfect Bayesian Equilibrium only requires the existence of supporting off-the-equilibrium belief. Sequential equilibrium additionally requires the supporting off-the- equilibrium belief to be consistent in the following sense: there *exists* a tremble in worker's strategy so that the off-the-equilibrium beliefs can be reached and updated with Bayes rule, and when the tremble goes to zero, the limit of the off-the-equilibrium beliefs need to converge to the supporting beliefs. Since in the separating equilibrium, both information sets of employers are reached, so the separating equilibrium is sequential equilibrium. Now check the consistency of the supporting off-the-equilibrium for the pooling PBE in which $e_h = e_l = 0$.

Let the high type choose $e_h = 1$ with probability ε^2, and the low type with ε. Then the off-the-equilibrium belief μ_1 can be updated as

$$\mu_1 = \frac{\mu\varepsilon(1-\lambda_l)}{\mu\varepsilon(1-\lambda_l)+(1-\mu)\varepsilon^2(1-\lambda_h)}$$

When ε goes to zero, μ_1 goes to 1, which is one of the off-the-equilibrium beliefs that supports the pooling strategy of workers. Hence, the pooling PBE is also a sequential equilibrium of this market.

[1] An implicit assumption in this argument is such information is socially valuable. The social value could be the technological innovations in the long run as a result of better match between workers' abilities and job types. Plus, when discharge of an employee is costly and the employers are risk averse, the elimination of employer's uncertainty on worker's type will promote hiring in the labor market.

APPENDIX D

INSTRUCTIONS FOR THE EXPERIMENTS IN CHAPTER 1

Instructions for Part I

Welcome and thank you for participating in this experiment. The experiment has 2 parts; this is the instruction for Part I. Please read the instructions carefully to help you understand the task and earn more money.

In this part, you will trade products in series of markets. In each market, three people will interact with each other: ***one person*** plays as the seller who is selling ***one*** product, and ***the other two people*** will be the buyers who are bidding for this product. The buyer who submits the higher bid will win the product; if the bids are equal, then one buyer will be randomly selected as the winner. A deal will be made between the seller and the winning buyer, at a price equal to the winning bid.

You will trade with other participants for 32 market periods in Part I; you will then proceed to Part II to make some more money before we conclude this experiment and pay your earnings. Instructions for Part II will be given after Part I is completed.

Groups and Roles

All participants will be *randomly* divided into small groups of 3 people in *each* period; each small group forms a market. One unit of product will be traded *within* each market; **you will only trade with the 2 partners in your own group.** The random grouping procedure will be performed before each period, so keep in mind that **you will NOT trade with the same partners in two different periods.** *IDs of group members will not be disclosed at any time.*

Each participant gets to play buyer in some periods, and seller in the other periods. Throughout these 32 periods, your role will change back and forth; generally, you will play consecutively as one role for a few periods before switching to the other. You'll find out your role as soon as a trading period begins; it will also be highlighted at the top section of your computer screen throughout that period.

Trading in a Market

There are two types of products with high or low quality; each high quality product has a value of $25, and each low quality product has a value of $10. After groups are formed, each seller will get one product to sell. *1/2* of the sellers will get the high quality products; *1/2* sellers will get the low quality ones. **The seller will be informed of the quality of the product she gets as soon as a period begins; the 2 buyers will NOT be informed of the quality of the product UNTIL the end of each period.**

The seller will then decide if she wants to spend $5 to take a quality test on her product: there is 90% chance for the high quality product to pass the test; 10% chance for the low quality product to pass. **Upon passing the test, a quality certificate will be issued to the product.** Notice that the $5 will be deducted from seller's profit if she decides to take the test, no matter if the product passes or fails.

Then the product will be placed on the market, and each buyer will need to submit a bid. The buyer with the higher bid wins the product. The buyers can submit the bid in the range of $0 to $30, in multiples of $0.25. **Before submitting their bids, the buyers will be informed if the product comes with a certificate, but NOT the seller's decision.** After both buyers submit the bids, they will be informed of the winner as well as the quality of the product in their market in that period**.**

The product will be sold to the winner at a price equal to his bid. Seller's profit is equal to this price (that is, the winning bid); ***minus*** the $5 cost of test if the seller decided to take the test. The winner of the product will get the product value ($25 or $10, depending on the product that he wins) ***minus*** his bid. ***In addition***, both buyers (that is, also the buyer who did not win the product) receives $8 in every period. Everyone will be informed of their earnings in the current periods before moving on to next period.

The profit of everyone in the market of each period is also shown in the following table:

Buyer with lower Bid	$8
Buyer with higher Bid	**$25** – Bid +$8 if the buyer wins a high quality product **$10** – Bid +$8 if the buyer wins a low quality product
Seller	Product Price – $5 if she took the test Product Price if she *didn't* take the test

The Quality Test

If you are a seller and decide to take the test, you will see a bar with a success region, a failure region and a randomly moving needle. The region of success is 90% of the bar if your product has high quality (Figure 1); the region of success is 10% of the bar if your product has low quality (Figure 2). The needle is equally likely to appear at any position along the bar, and it will stop moving after 5 seconds. If the needle ends up in the success region, your product passes the test and will get a quality certificate; if it ends up in the failure region, your product fails the test and will be placed on the market without a quality certificate.

Figure 1 High Quality Product Test

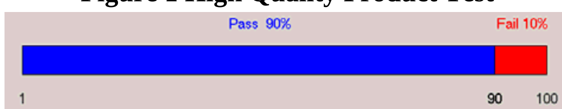
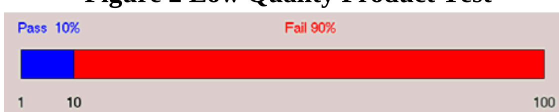


Figure 2 Low Quality Product Test



Summary: a trading period proceeds as follows

⇒ Markets are formed randomly
⇒ Products are distributed to sellers randomly

Then within each market:

⇒ Seller learns the value of her product
⇒ Seller decides whether to take the test
⇒ Product is placed on the market, with or without a quality certificate
⇒ Buyers make bids
⇒ Buyers learn the quality of the product in their market; Seller and buyers learn results and profits

Payment

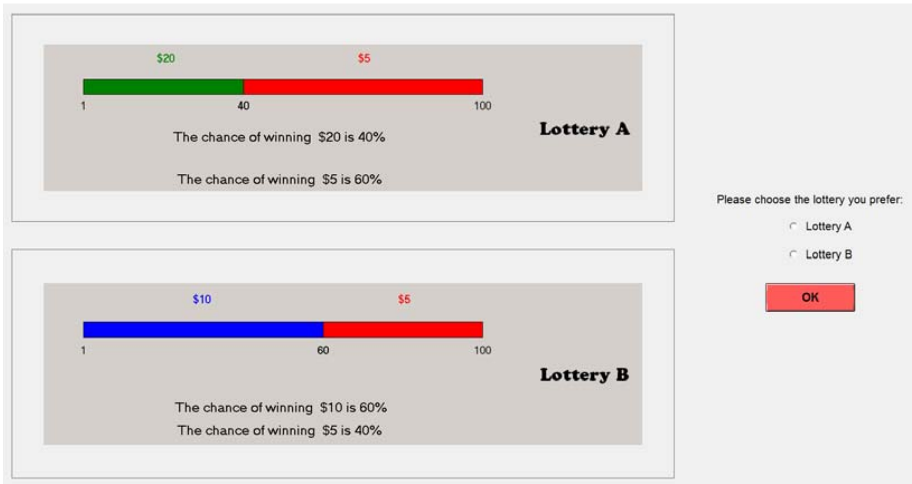
After Part I is completed, one of these 32 market periods will be randomly chosen by drawing a numbered ball from the bingo cage, and your payoff in the chosen period will be your earnings from Part I. Since you don't know which period will be chosen, please decide carefully in every period.

Instructions for Part II

In this part of the experiment, you will see pairs of lotteries. Your task is to choose which lottery you prefer to play in each pair. ***After you make each choice, the lottery of your choice will be played out to determine your payoff in that task. There is no right or wrong answer, simply pick the one you prefer to play.***

Each lottery will offer monetary prizes with some probabilities. Different lotteries might differ in prizes and probabilities. You will see one pair of lotteries at a time, and here is an example of what you will see on your computer screen:

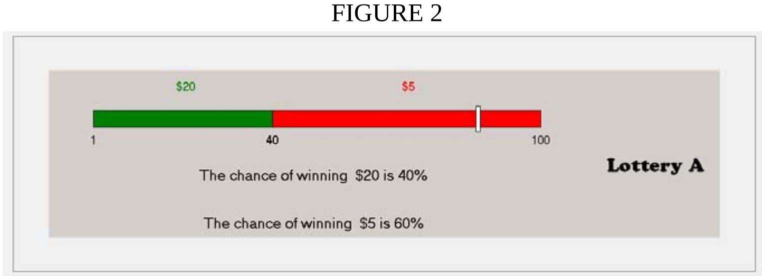
FIGURE 1



You can tell the chances of winning each prize from the length of same-colored bar and the descriptions below the bar. In Figure 1, **Lottery A** offers $20 with 40% chance, so length of green bar is 40% of the whole bar; it offers $5 with 60% chance, so the red bar is 60% length of the whole bar. **Lottery B** offers a prize of $10 with a 60% chance, so the length of the blue bar is 60% length of the whole bar; it offers a prize of $5 with a 40% chance, so the length of the red bar is the other 40% length of the whole bar. Notice that different lotteries might have these colors matched to different prizes, for example if a lottery involves prizes $20, $15 and $5, they will also be shown in green, blue, red, respectively.

You can choose your preferred lottery by clicking on the corresponding button on the right. After you click OK, you will see your chosen lottery on the screen and a white needle will show up. The needle is equally likely to appear at any position along the bar, and it will stop moving after 5 seconds. The position of the needle ends up in will be the prize you win, and it will be your payoff of this lottery task. For example, if you have chosen Lottery A and the needle in Figure 2 stopped in the red region with $5 on top of it, so the payoff is $5.

FIGURE 2



After you make all 20 choices, 20 balls numbered 1 to 20 will be put in the bingo cage; a numbered ball will be randomly drawn, and your payoff in that lottery task will be your earnings in this part. ***Please choose carefully in each task, as you don't know which task will be chosen until after all 20 decisions are made.***

APPENDIX E

ESTIMATION UNDER THE SWITCHING ASSUMPTION

In this appendix we derive the posterior distribution under the assumption that subjects switch decision modes over the 100 lottery tasks, and present estimation results using the algorithms developed in Section 2.4. Under this switching assumption, the mixture probability ρ should be interpreted as the proportion of *choices* that can be characterized as the EUT type, rather than the proportion of *subjects*. We keep the assumption that there is no individual heterogeneity among subjects: subjects switch between the EUT and CPT modes with the same the mixture probability and have the same preference parameters when using each decision mode.

We simulate a different dataset of 177 subjects' choices over the same 100 lottery pairs and using the same values of parameters as in Section 2.2 under the switching assumption. We summarize the mixture probability of the EUT type choices and preference parameters for both decision types as follows:

Percentage of EUT Choices: $\rho = 0.6$

EUT CRRA: $r = 0.3$

EUT Noise: $\mu_{EUT} = 2$

CPT Gain-Domain CRRA: $\alpha = 0.1$

CPT Loss-Domain CRRA: $\beta = 0.5$

CPT Loss Aversion Parameter: $\lambda = 1.8$

CPT Gain-Domain Weighting Function: $\gamma^{+} = 0.8$

CPT Gain-Domain Weighting Function: $\gamma^{-} = 0.6$

CPT Noise: $\mu_{CPT} = 2$

Also assuming that the behavioral noises are identically and independently distributed for all subjects and decisions, the likelihood of an arbitrary subject i's choices y_i is

$$p(y_i|\theta) = \prod_{t=1}^{100} \left[\rho G(\nabla EU_{it})^{y_{it}} \cdot (1 - G(\nabla EU_{it}))^{1-y_{it}} + (1-\rho) G(\nabla CPU_{it})^{y_{it}} \cdot (1 - G(\nabla CPU_{it}))^{1-y_{it}} \right]$$

The likelihood for observing all subjects' decisions $y = \{y_1, \cdots, y_{177}\}$ is

$$p(y|\theta) = \prod_{i=1}^{177} \prod_{t=1}^{100} \left[\rho G(\nabla EU_{it})^{y_{it}} \cdot (1 - G(\nabla EU_{it}))^{1-y_{it}} + (1-\rho) G(\nabla CPU_{it})^{y_{it}} \cdot (1 - G(\nabla CPU_{it}))^{1-y_{it}} \right]$$

We assume parameters are independently distributed in the prior distribution and use the same non-informative marginal priors as in Equation 2.3 of Section 2.4. Combining prior with likelihood we have the posterior distribution under the switching assumption $p(\theta|y) \propto p(y|\theta)p(\theta)$.

The MCMC algorithms to sample from this posterior distribution are the same with those in Section 2.4 in terms of structures. For the algorithms *without* data augmentation for allocations in

Section 2.4.1, we only need to substitute the new posterior distribution derived above in evaluating the acceptance probabilities. For the algorithms *with* data augmentation for allocations in Section 2.4.2, we derive the posterior distributions of allocations, preference parameters and mixture probability below. Under the assumption that subjects switch modes across decisions, this algorithm will require allocating *each choice y_{it}* to either EUT or CPT mode.

With some initial value of $\theta^{(0)} = \{\rho^{(0)}, r^{(0)}, \mu_{EUT}^{(0)}, \alpha^{(0)}, \beta^{(0)}, \lambda^{(0)}, \gamma^{+(0)}, \gamma^{-(0)}, \mu_{EUT}^{(0)}\}$, the m^{th} iteration is:

(1) Sample allocation $S_{it}^{(m)}$ from $p(S_{it}|y_{it}, \theta^{(m-1)})$:

$$
\begin{aligned}
Pr(S_{it} = 1|y_{it}, \theta^{(m-1)}) &\propto p(y_{it}|S_{it} = 1, \theta^{(m-1)}) Pr(S_{it} = 1|\theta^{(m-1)}) \\
&= p(y_{it}|\theta_{EUT}^{(m-1)})\rho^{(m-1)} \\
Pr(S_{it} = 0|y_{it}, \theta^{(m-1)}) &\propto p(y_{it}|S_{it} = 0, \theta^{(m-1)}) Pr(S_{it} = 0|\theta^{(m-1)}) \\
&= p(y_{it}|\theta_{CPT}^{(m-1)})(1 - \rho^{(m-1)}) \\
S_{it} &\sim \text{Bernoulli}\left(\frac{Pr(S_{it} = 1|y_{it}, \theta^{(m-1)})}{Pr(S_{it} = 1|y_{it}, \theta^{(m-1)}) + Pr(S_{it} = 0|y_{it}, \theta^{(m-1)})} \right)
\end{aligned}
$$

Collect the allocations for all the choices in $S^{(m)}$.

(2) Sample parameters $\theta^{(m)}$ from $p(\theta^{(m)}|y, S^{(m)})$:

$$
\begin{aligned}
p(y, S|\theta) &= \prod_{i=1}^{177} p(y_{it}|S_{it}, \theta) p(S_{it}|\theta) \\
&= \prod_{\{i,t\}:S_{it}=1} \Big(\rho \cdot p(y_{it}|\theta_{EUT}) \Big) \times \prod_{\{i,t\}:S_{it}=0} \Big((1-\rho) \cdot p(y_{it}|\theta_{CPT}) \Big) \\
&= \underbrace{\prod_{\{i,t\}:S_{it}=1} p(y_{it}|\theta_{EUT})}_{\text{Likelihood of EUT Choices}} \times \underbrace{\prod_{\{i,t\}:S_{it}=0} p(y_{it}|\theta_{CPT})}_{\text{Likelihood of CPT Choices}} \times \underbrace{\rho^{N_1}(1-\rho)^{N_0}}_{\text{Likelihood of Allocations}}
\end{aligned}
$$

The superscripts for iteration are also dropped here. N_1 is the number of *choices* classified as EUT, and N_0 is the number of *choices* classified as CPT from step (1). Under the assumption that parameters are independently distributed in their prior distribution, we have

$$p(\theta|y,S) \propto p(y,S|\theta)p(\theta)$$
$$= \underbrace{\prod_{\{i,t\}:S_{it}=1} p(y_{it}|\theta_{EUT}) \cdot p(\theta_{EUT})}_{\text{Posterior of EUT Choices}} \times \underbrace{\prod_{\{i,t\}:S_{it}=0} p(y_{it}|\theta_{CPT}) \cdot p(\theta_{CPT})}_{\text{Posterior of CPT Choices}} \times \underbrace{\rho^{N_1}(1-\rho)^{N_0} \cdot p(\rho)}_{\text{Posterior of Allocations}}$$

Then, in step (2) of the m^{th} iteration, we can separately sample θ_{EUT}, θ_{CPT} and ρ from their respective posteriors as in Section 2.4.2.

In contrast to the performance of these algorithms under the no-switching assumption, the algorithms generally fail to converge under the switching assumption with only one exception, the algorithm with no data augmentation for allocations and with the parameters individually updated (the third variant in Section 2.4.1). For the posterior samples with this algorithm, we report the trace plots in Figure E.1, the histograms after burn-in in Figure E.2 and the autocorrelation plots in Figure E.3. Generally, the MCMC samples have higher autocorrelations and lower efficiencies compared to those under the no-switching assumption. We report summary statistics in Table E.1. We find that the true parameters are generally within the 90% confidence intervals, despite that the loss-domain probability weighting parameter γ^- is very close to the upper bound.

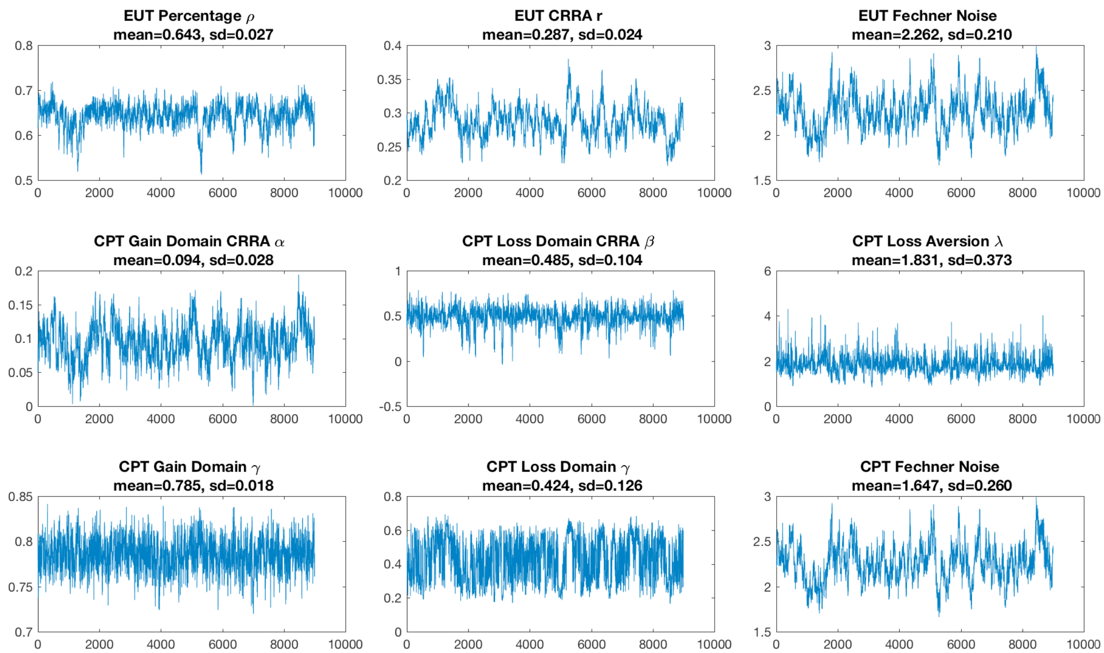
We also apply the algorithms with the posterior distribution under switching assumption to the analysis of observed choice data. Again, we only observe convergence with the algorithm without label augmentation and with the parameters individually updated converges. We report summary statistics in Table E.2. We provide the trace plots in Figure E.4, the histograms in Figure E.5 and the autocorrelation plots in Figure E.6. Under the switching assumption, we find that 29.5% of all choices can be characterized as the EUT type, while under the no-switching assumption, 56.5% of all subjects can be characterized as the EUT type. We find higher levels of risk aversion among the EUT type choices, in contrast to the moderate risk aversion among the EUT type subjects under the no-switching assumption. Among the CPT type choices, we find virtually no evidence of loss aversion, and find over-estimation of probabilities with extreme outcomes only in the gain-domain.

Table E.1: Estimation Results Using Simulated Subject Choices under Switching Assumption

Paramters	**True Values**	**Metropolis Hasting Without Label Augmentation**			
		Mean	**S.D.**	**5% Quantile**	**95% Quantile**
ρ	**0.6**	0.643	0.027	0.592	0.682
r	**0.3**	0.287	0.024	0.249	0.328
μ_{EUT}	**2.0**	2.263	0.210	1.915	2.626
α	**0.1**	0.094	0.028	0.048	0.141
β	**0.5**	0.486	0.104	0.297	0.633
λ	**1.8**	1.831	0.373	1.249	2.494
γ^{+}	**0.8**	0.785	0.018	0.756	0.814
γ^{-}	**0.6**	0.424	0.126	0.225	0.608
μ_{CPT}	**2.0**	1.647	0.260	1.229	2.105

Table E.2: Estimation Results Using Observed Subject Choices under Switching Assumption

Paramters	**Mean**	**S.D.**	**5% Quantile**	**95% Quantile**
ρ	0.295	0.011	0.277	0.314
r	0.854	0.001	0.852	0.856
μ_{EUT}	0.074	0.009	0.062	0.090
α	0.113	0.023	0.072	0.151
β	0.145	0.022	0.109	0.181
λ	1.038	0.034	0.985	1.096
γ^{+}	0.730	0.014	0.706	0.754
γ^{-}	0.949	0.036	0.893	1.010
μ_{CPT}	4.031	0.327	3.533	4.606



[1] In Figures E.1, E.2 and E.3: Observed subjects' choices from Harrison and Swarthout (2016) are used.
[2] In Figures E.1, E.2 and E.3: Assume subjects switch between EUT and CPT.
[3] The mean and standard deviation for each sample are noted above each panel.

Figure E.1: Trace Plots of Metropolis Hastings Samples Based on Simulated Observations under Switching Assumption

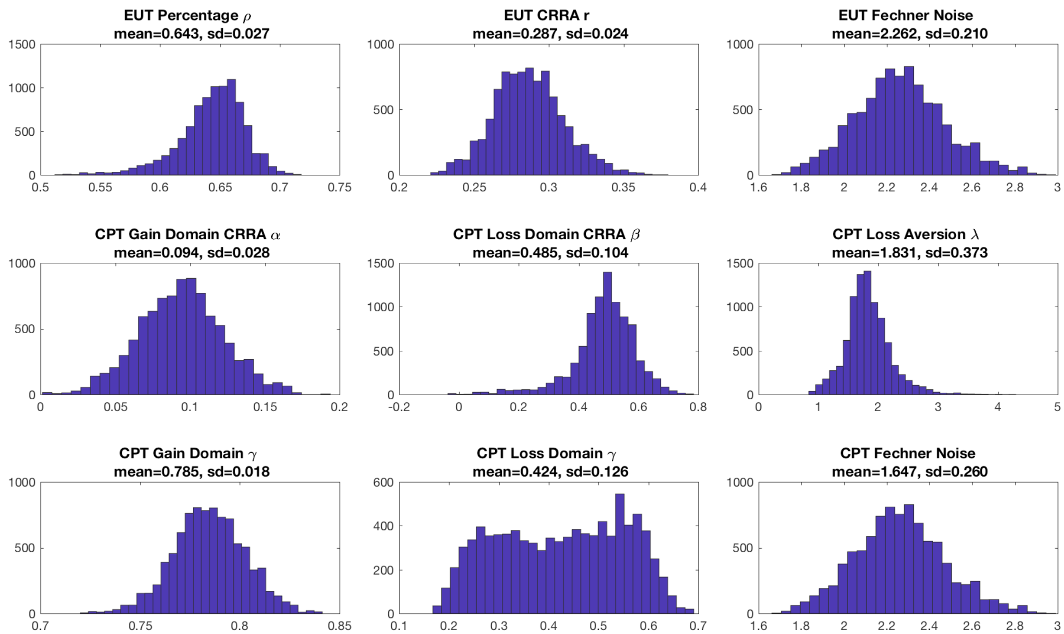


Figure E.2: Histograms of Metropolis Hastings Samples Based on Simulated Observations under Switching Assumption

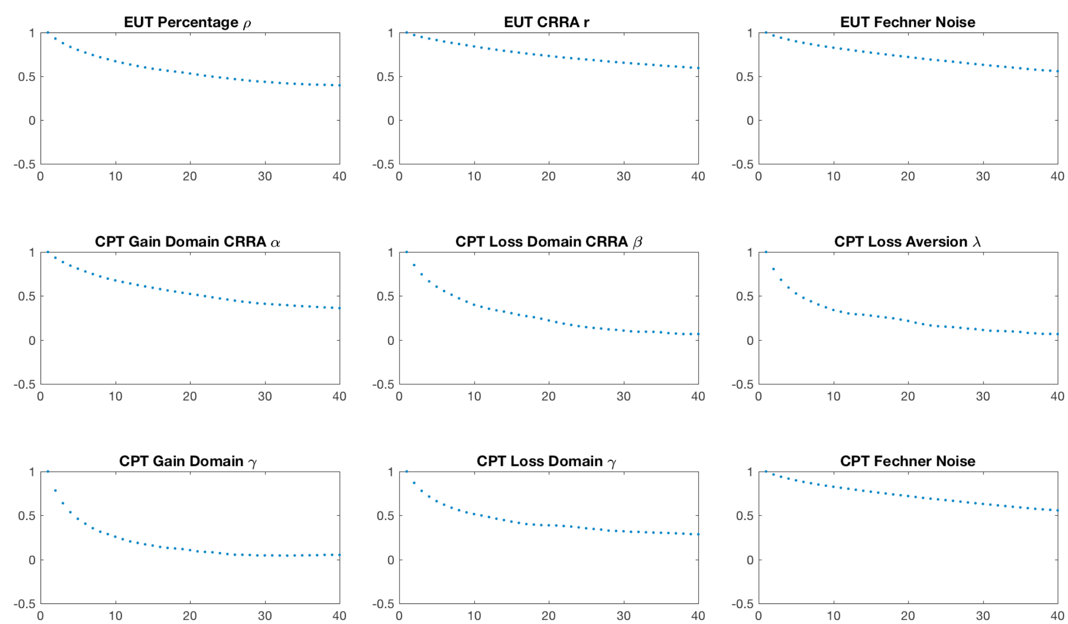
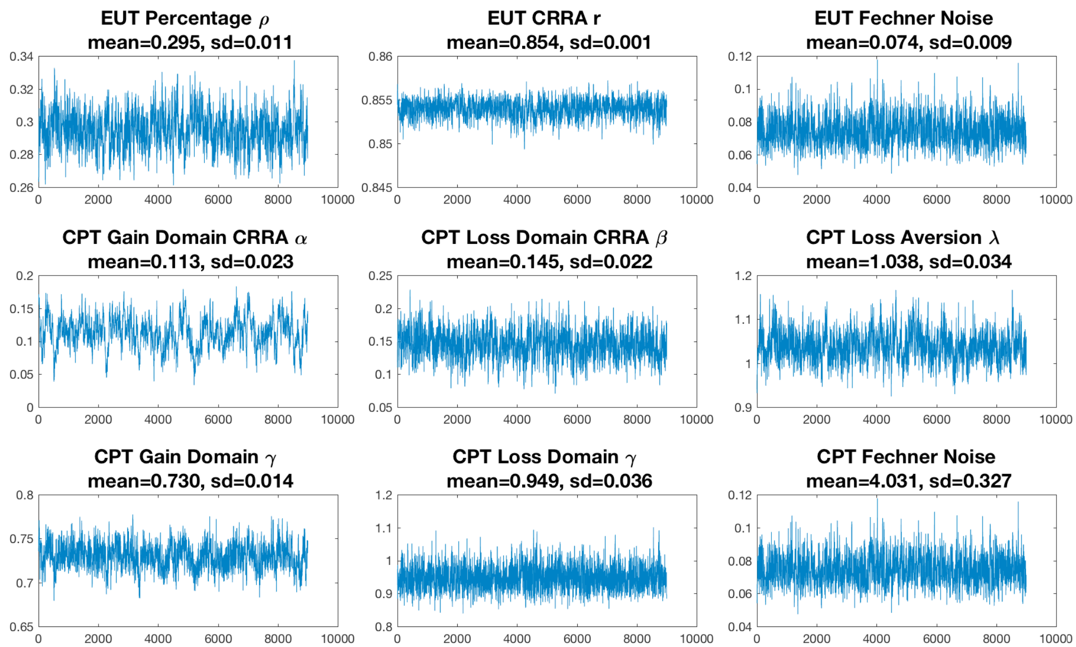


Figure E.3: Autocorrelation of MH Sample Based on Simulated Observations under Switching Assumption



[1] In Figures E.4, E.5 and E.6: Observed subjects' choices from Harrison and Swarthout (2016) are used.
[2] In Figures E.4, E.5 and E.6:Assume subjects switch between EUT and CPT.
[3] The mean and standard deviation for each sample are noted above each panel.

Figure E.4: Trace Plots of Metropolis Hastings Samples Using Observed Subject Choices under Switching Assumption

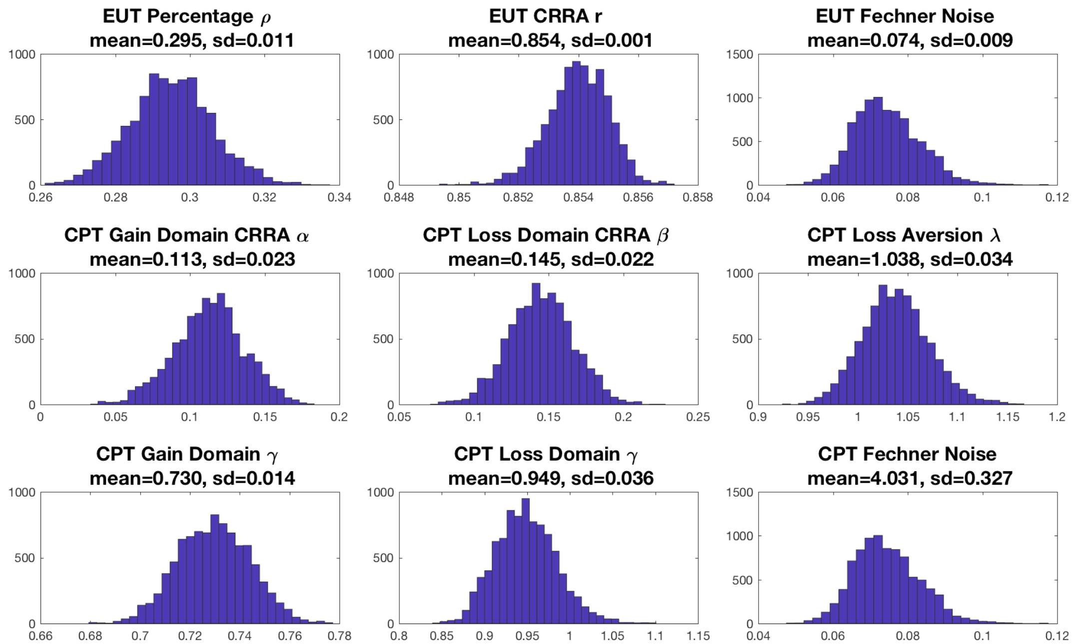


Figure E.5: Histograms of Metropolis Hastings Samples Using Observed Subject Choices under Switching Assumption

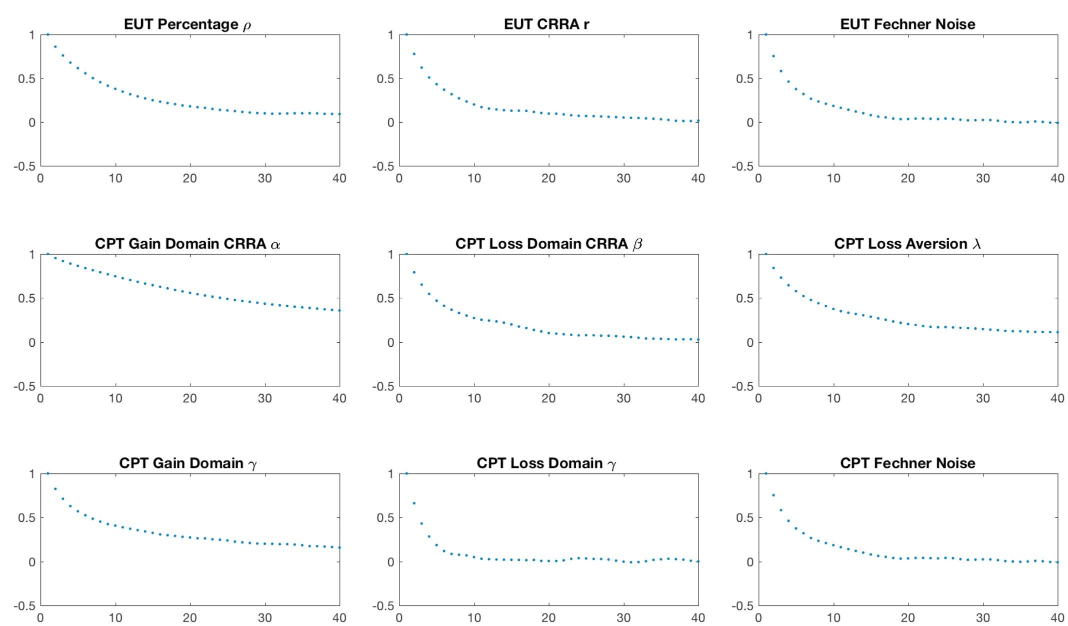


Figure E.6: Autocorrelation Plots of Metropolis Hastings Samples Using Observed Subject Choices under Switching Assumption

Bibliography

Abdellaoui, M., H. Bleichrodt, and C. Paraschiv (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science 53*(10), 1659–1674.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2008). Eliciting risk and time preferences. *Econometrica 76*(3), 583–618.

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

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2014b). Dual criteria decisions. *Journal of Economic Psychology 41*, 101–113.

Banks, J., C. Camerer, and D. Porter (1994). An experimental analysis of nash refinements in signaling games. *Games and Economic Behavior 6*(1), 1–31.

Bardsley, N. and P. G. Moffatt (2007). The experimetrics of public goods: Inferring motivations from contributions. *Theory and Decision 62*(2), 161–193.

Becker, G. M., M. H. DeGroot, and J. Marschak (1963). Stochastic models of choice behavior. *Behavioral Science 8*(1), 41–55.

Belman, D. and J. S. Heywood (1991). Sheepskin effects in the returns to education: An examination of women and minorities. *The Review of Economics and Statistics 73*(4), 720–724.

Brandts, J. and C. A. Holt (1992). An experimental test of equilibrium dominance in signaling games. *The American Economic Review 82*(5), 1350–1365.

Brandts, J. and C. A. Holt (1993). Adjustment patterns and equilibrium selection in experimental signaling games. *International Journal of Game Theory 22*(3), 279–302.

Bruhin, A., H. Fehr-Duda, and T. Epper (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica 78*(4), 1375–1412.

Bueno, H., J. S. Ross, Y. Wang, J. Chen, M. T. Vidán, S.-L. T. Normand, J. P. Curtis, E. E. Drye, J. H. Lichtman, P. S. Keenan, et al. (2010). Trends in length of stay and short-term outcomes among medicare patients hospitalized for heart failure, 1993-2006. *The Journal of the American Medical Association 303*(21), 2141–2147.

Burnham, K. P. and D. R. Anderson (2003). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media.

Cadsby, C. B., M. Frank, and V. Maksimovic (1990). Pooling, separating, and semiseparating equilibria in financial markets: Some experimental evidence. *Review of Financial Studies 3*(3), 315–342.

Cadsby, C. B., M. Frank, and V. Maksimovic (1998). Equilibrium dominance in experimental financial markets. *Review of Financial Studies 11*(1), 189–232.

Carlsson, H. and S. Dasgupta (1997). Noise-proof equilibria in two-action signaling games. *Journal of Economic Theory 77*(2), 432–460.

Centers for Medicare & Medicaid Services (2007). Medicare & medicaid statistical supplement. *Baltimore*.

Charness, G., U. Gneezy, and A. Imas (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization 87*, 43–51.

Cho, I.-K. and D. M. Kreps (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics 102*(2), 179–221.

Conlisk, J. (1989). Three variants on the allais example. *American Economic Review 79*(3), 392–407.

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

Cooper, D. J., S. Garvin, and J. H. Kagel (1997). Signalling and adaptive learning in an entry limit pricing game. *The RAND Journal of Economics*, 662–683.

Cooper, D. J. and J. H. Kagel (2003). The impact of meaningful context on strategic play in signaling games. *Journal of Economic Behavior & Organization 50*(3), 311–337.

Cox, J. C. and V. Sadiraj (2006). Small-and large-stakes risk aversion: Implications of concavity calibration for decision theory. *Games and Economic Behavior 56*(1), 45–60.

Cox, J. C., V. Sadiraj, and U. Schmidt (2015). Paradoxes and mechanisms for choice under risk. *Experimental Economics 18*(2), 215–250.

Dawes, A. J., G. D. Sacks, M. M. Russell, A. Y. Lin, M. Maggard-Gibbons, D. Winograd, H. R. Chung, A. Tillou, J. R. Hiatt, and C. Ko (2014). Preventable readmissions to surgical services: Lessons learned and targets for improvement. *Journal of the American College of Surgeons 219*(3), 382 – 389.

de Haan, T., T. Offerman, and R. Sloof (2011). Noisy signaling: Theory and experiment. *Games and Economic Behavior 73*(2), 402–428.

El-Gamal, M. A. and D. M. Grether (1995). Are people bayesian? Uncovering behavioral strategies. *Journal of the American Statistical Association 90*(432), 1137–1145.

Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer Science & Business Media.

Gelman, A., G. O. Roberts, and W. R. Gilks (1996). Efficient metropolis jumping rules. *Bayesian Statistics 5*, 599–607.

Glance, L. G., A. L. Kellermann, T. M. Osler, Y. Li, D. B. Mukamel, S. J. Lustik, M. P. Eaton, and A. W. Dick (2014). Hospital readmission after noncardiac surgery: the role of major complications. *The Journal of the American Medical Association surgery 149*(5), 439–445.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica 59*(3), 667.

Halverson, A. L., M. M. Sellers, K. Y. Bilimoria, M. T. Hawn, M. V. Williams, R. S. McLeod, and C. Y. Ko (2014). Identification of process measures to reduce postoperative readmission. *Journal of Gastrointestinal Surgery 18*(8), 1407–1415.

Harrison, G. W., E. Johnson, M. M. McInnes, and E. E. Rutström (2007). Measurement with experimental controls. In M. Boumans (Ed.), *Measurement in Economics: A Handbook*. San Diego, CA: Elsevier.

Harrison, G. W. and J. M. Ng (2016). Evaluating the expected welfare gain from insurance. *Journal of Risk and Insurance 83*(1), 91–120.

Harrison, G. W. and E. E. Rutström (2008). Risk aversion in the laboratory. In J. C. Cox and G. W. Harrison (Eds.), *Risk Aversion in Experiments*, pp. 41–196. Bingley: Emerald.

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

Harrison, G. W. and J. T. Swarthout (2014). Experimental payment protocols and the bipolar behaviorist. *Theory and Decision 77*(3), 423–438.

Harrison, G. W. and J. T. Swarthout (2016). Cumulative prospect theory in the laboratory: A reconsideration. *CEAR Working Paper 2016-05.*. Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, 2016.

Haruvy, E., D. O. Stahl, and P. W. Wilson (2001). Modeling and testing for heterogeneity in observed strategic behavior. *Review of Economics and Statistics 83*(1), 146–157.

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

Heywood, J. S. (1994). How widespread are sheepskin returns to education in the u.s.? *Economics of Education Review 13*(3), 227 – 234.

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

Horwitz, L., C. Partovian, Z. Lin, J. Herrin, J. Grady, M. Conover, J. Montague, C. Dillaway, K. Bartczak, J. Ross, et al. (2011). Hospital-wide (all-condition) 30-day risk-standardized readmission measure. *Yale New Haven Health Services Corporation/Center for Outcomes Research & Evaluation. Retrieved September 10*, 2012.

Houser, D., M. Keane, and K. McCabe (2004). Behavior in a dynamic decision problem: An analysis of experimental evidence using a bayesian type classification algorithm. *Econometrica 72*(3), 781–822.

Hungerford, T. and G. Solon (1987). Sheepskin effects in the returns to education. *The Review of Economics and Statistics 69*(1), 175–177.

Hurley, T. M. and J. F. Shogren (2005). An experimental comparison of induced and elicited beliefs. *Journal of Risk and Uncertainty 30*(2), 169–188.

Jaeger, D. a. and M. E. Page (1996). Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education. *The Review of Economics and Statistics 78*(4), 733–740.

Jeitschko, T. D. and H.-T. Normann (2012). Signaling in deterministic and stochastic settings. *Journal of Economic Behavior & Organization 82*(1), 39–55.

Jencks, S. F., M. V. Williams, and E. A. Coleman (2009). Rehospitalizations among patients in the medicare fee-for-service program. *New England Journal of Medicine 360*(14), 1418–1428. PMID: 19339721.

Joynt, K. E. and A. K. Jha (2013). Characteristics of hospitals receiving penalties under the hospital readmissions reduction program. *The Journal of the American Medical Association 309*(4), 342–343.

Kaboli, P. J., J. T. Go, J. Hockenberry, J. M. Glasgow, S. R. Johnson, G. E. Rosenthal, M. P. Jones, and M. Vaughan-Sarrazin (2012). Associations between reduced hospital length of stay and 30-day readmission rate and mortality: 14-year experience in 129 veterans affairs hospitals. *Annals of Internal Medicine 157*(12), 837–845.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Kansagara, D., H. Englander, A. Salanitro, D. Kagen, C. Theobald, M. Freeman, and S. Kripalani (2011). Risk prediction models for hospital readmission: a systematic review. *The Journal of the American Medical Association 306*(15), 1688–1698.

Kassin, M. T., R. M. Owen, S. D. Perez, I. Leeds, J. C. Cox, K. Schnier, V. Sadiraj, and J. F. Sweeney (2012). Risk factors for 30-day hospital readmission among general surgery patients. *Journal of the American College of Surgeons 215*(3), 322–330.

Kohlnhofer, B. M., S. E. Tevis, S. M. Weber, and G. D. Kennedy (2014). Multiple complications and short length of stay are associated with postoperative readmissions. *The American Journal of Surgery 207*(4), 449–456.

Kübler, D., W. Müller, and H.-T. Normann (2008). Job-market signaling and screening: An experimental comparison. *Games and Economic Behavior 64*(1), 219–236.

Landeras, P. and J. Villarreal (2005). A noisy screening model of education. *Labour 19*(1), 35–54.

Lawson, E. H., B. L. Hall, R. Louie, S. L. Ettner, D. S. Zingmond, L. Han, M. Rapp, and C. Y. Ko (2013). Association between occurrence of a postoperative complication and readmission: implications for quality improvement and cost savings. *Annals of surgery 258*(1), 10–18.

Leeds, I. L., V. Sadiraj, J. C. Cox, K. E. Schnier, and J. F. Sweeney (2013). Assessing clinical discharge data preferences among practicing surgeons. *Journal of Surgical Research 184*(1), 42–48.

Loomes, G. and R. Sugden (1995). Incorporating a stochastic element into decision theories. *European Economic Review 39*(3-4), 641–648.

Loomes, G. and R. Sugden (1998). Testing different stochastic specifications of risky choice. *Economica 65*, 581–598.

Lucas, D. J., A. Haider, E. Haut, R. Dodson, C. L. Wolfgang, N. Ahuja, J. Sweeney, and T. M. Pawlik (2013). Assessing readmission after general, vascular, and thoracic surgery using acs-nsqip. *Annals of Surgery 258*(3), 430.

Matthews, S. A. and L. J. Mirman (1983). Equilibrium limit pricing: The effects of private information and stochastic demand. *Econometrica 51*(4), pp. 981–996.

McLachlan, G. and D. Peel (2000). *Finite Mixture Models*. New York: Wiley.

Merkow, R. P., M. H. Ju, J. W. Chung, B. L. Hall, M. E. Cohen, M. V. Williams, T. C. Tsai, C. Y. Ko, and K. Y. Bilimoria (2015). Underlying reasons associated with hospital readmission following surgery in the united states. *The Journal of the American Medical Association 313*(5), 483–495.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics 21*(6), 1087–1092.

Miller, R. M. and C. R. Plott (1985). Product quality signaling in experimental markets. *Econometrica 83*(4), 837–872.

Morris, M. S., R. J. Deierhoi, J. S. Richman, L. K. Altom, and M. T. Hawn (2014). The relationship between timing of surgical complications and hospital readmission. *The Journal of the American Medical Association surgery 149*(4), 348–354.

Murphy, R. O. and R. H. W. ten Brincke (2017). Hierarchical maximum likelihood parameter estimation for cumulative prospect theory: Improving the reliability of individual risk parameter estimates. *Management Science Advance online publication. doi:10.1037/a0028240*.

Neumann, J. V. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Nilsson, H., J. Rieskamp, and E.-J. Wagenmakers (2011). Hierarchical bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology 55*(1), 84–93.

of the Patient Protection, C. and A. C. Act (2010). Office of legislative counsel. *U.S. House of Representatives, United States of America*.

Park, J. H. (1999). Estimation of sheepskin effects using the old and the new measures of educational attainment in the Current Population Survey. *Economics Letters 62*, 237–240.

Posey, L. L. and A. Yavas (2007). Screening equilibria in experimental markets. *The Geneva Risk and Insurance Review 32*(2), 147–167.

Potters, J. and F. van Winden (1996). Comparative statics of a signaling game: An experimental study. *International Journal of Game Theory 25*(3), 329–353.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization 3*(4), 323–343.

Regev, T. (2012, August). Education signaling with uncertain returns. *The B.E. Journal of Theoretical Economics 12*(1), 1935–1704.

Schoemaker, P. J. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature 20*(2), 529–563.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics 87*(3), 355–374.

Stahl, D. O. and P. W. Wilson (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior 10*(1), 218–254.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature 38*(2), 332–382.

Tsai, T. C., K. E. Joynt, E. J. Orav, A. A. Gawande, and A. K. Jha (2013). Variation in surgical-readmission rates and quality of hospital care. *New England Journal of Medicine 369*(12), 1134–1142. PMID: 24047062.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

Wakker, P. P. (2010). *Prospect Theory: For Risk and Ambiguity*. Cambridge University Press.

Wilcox, N. T. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. C. Cox and G. W. Harrison (Eds.), *Risk Aversion in Experiments*, pp. 197–292. Bingley: Emerald.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica 55*(1), 95.

Vita

Xiaoxue (Sherry) Gao was born on July 27, 1985 in Shandong, China. Sherry graduated from the double major program at Zhongnan University of Economics and Law in 2007 with a B.A. in Economics and a B.A. in Law. In 2009, Sherry graduated from Renmin University of China with a M.S. in Economics. Later in 2009, Sherry started her education in US as a master student in Statistics, and in 2011, as a Ph.D student in Economics in Georgia State University. Sherry's research interests focus on applied game theory, decision making under risk and uncertainty, behavioral and experimental economics and applied Bayesian Econometrics.

Most of her time as a Ph.D. student at Georgia State University, Sherry worked as a research and teaching assistant for Dr. Vjollca Sadiraj. She also worked as assistant for Dr. James C. Cox in oversea workshops of experimental economics in the summers of 2015 and 2016. In addition, Sherry also independently taught undergraduate courses, including The Global Economy and Introduction to Game Theory. Sherry actively participated in academic activities and presented her work at the Southern Economic Association, Economic Science Association. Sherry received the Andrew Young School Dissertation Fellowship, Provost's Fellowship and was awarded as scholar of Center for the Economic Analysis of Risks.

Sherry was awarded a Ph.D. in Economics by Georgia State University in August, 2017. She begins working as an Assistant Professor in Department of Resource Economics, College of Social and Behavioral Science at University of Massachusetts Amherst.