Fall 2007

Essays on Risk and Incentives

Russell Engel
Sacred Heart University, engelr@sacredheart.edu

Recommended Citation

Engel, Russell P. "Essays on Risk and Incentives." Diss. Florida State University, 2007. Web.

7-17-2007

Essays on Risk and Incentives

Russell Paul Engel
Florida State University

THE FLORIDA STATE UNIVERSITY

COLLEGE OF SOCIAL SCIENCES

ESSAYS ON RISK AND INCENTIVES

By

RUSSELL PAUL ENGEL

A Dissertation submitted to the
Department of Economics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Fall Semester, 2007

The members of the Committee approve the Dissertation of Russell P. Engel defended on July 17, 2007.

R. Mark Isaac
Professor Directing Dissertation

Doug Stevens
Outside Committee Member

Timothy Salmon
Committee Member

David Macpherson
Committee Member

The Office of Graduate Studies has verified and approved the above named committee members.

To my wife Judy, and my grandfather Jim

ACKNOWLEDGEMENTS

I would first like to acknowledge my advisors Mark Isaac and Tim Salmon. I consider myself extremely lucky to have been their student, and I can't imagine better mentors.

I would also like to thank the John and Hallie Quinn Research Fellowship for research support during my time at Florida State University.

On a personal note, I would like to thank my grandparents (Jim and Carol Engel, and William and Helena Franco), and my parents (Cheryl Engel and William Franco) for always being supportive. Since I was raised by my grandparents, I always considered their younger children as siblings, so I would like to include them in a list of my brothers and sisters. Thanks to Jimmy, Shirley, Jenna, Corrie, and Summer for always giving me someone to talk to (not quite yet in Summer's case). My extended family played an important part in my life, so thanks to the: Engels, Francos, Coelhos, Morrises, Albergarias, Bordens, and Nguyens. Special thanks to Bill Morris for advice before my departure to Florida State, and Mike Bertoncini for allowing me to keep my street cred.

Lastly I would like to thank my wife Judy. She is beautiful, talented, and a harder worker than I have ever been. I would not have made it through this process without her. She is also one hell of a cook.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

I use economic experiments to investigate individual behavior under uncertainty. The first essay examines the consistency of risk preferences over two institutions. The two institutions I use are the first price sealed bid auction and a Holt-Laury lottery. There is some controversy as to whether or not observed overbidding in first price auction is actually caused by risk aversion or simply consistent with it. Behavior in the Holt-Laury lottery being caused by risk aversion is not in dispute. By having the same subjects participate in both institutions, I show that subjects' risk preferences in the lottery are consistent with subjects' risk preferences in the auction. This supports the notion that behavior in the first price sealed bid auction is in fact driven by risk aversion. I also find support for the constant relative risk aversion model (CRRAM). Using CRRAM I find that the risk parameters derived in the lottery are consistent with the risk parameters derived in the auction. The second essay examines how individuals respond to random monitoring. I design a real effort laboratory experiment with incentives similar to those faced by many workers. Subjects are allowed to engage two tasks; one task mimics work for an employer, the other task allows for gains due to shirking. Employee shirking has the potential to be extremely costly to firms. To counter the productivity loss due to shirking, firms may institute various monitoring schemes. Previous experimental research has shown that while monitoring does decrease shirking, some subjects work without explicit financial incentives. My experimental design corrects for the possibility of these past results being artifacts of past experimnetal design. I find that subjects who are given incentives to shirk do in fact shirk, but monitoring and an attainable quota lead to increased productivity. However, when the quota is unattainable, subjects revolt and engage in a high amount of shirking.

CHAPTER 1

Introduction

Economic models are built by making certain assumptions about individual preferences. When an individual will have to make a decision under uncertainty his risk preferences have to be part of the model. Risk preferences of individuals describe how they view trade-offs for risky propositions and are an integral part of a wide range of economic models (constant absolute risk aversion, decreasing absolute risk aversion, constant relative risk aversion, increasing relative risk aversion, hyperbolic absolute risk aversion, etc...). An important assumption is that an individuals risk preferences are constant across various situations. This allows for great convenience in modelling behavior across tasks. Should such parameters be task dependent then much additional study will be required to understand such behavioral shifts. Recent findings have shown that risk attitudes may not be consistent across institutions. If these past findings are robust, we are in the more difficult situation.

I use economic experiments to explore this issue. I place subjects in two settings that each allow for the elicitation of their risk preferences. The first institution is a lottery, and the second institution is an auction. There is an ongoing debate about whether or not the risk preferences derived from auctions are truly measuring risk or just some other behavior that looks like a risk preference. This means I cannot *prove* inconsistency of risk preferences. I can only show that the modelling of risk in both institutions is inconsistent. But if I do find that the parameters are consistent, not only are people consistent, but their behavior in auctions can, in fact, be explained by risk preferences. For the institutions that I examine, I will show that standard assumptions of consistent preferences do quite well. Individuals that are risk averse in one institution will be risk averse in the other institution. Furthermore, I show that a specific model is quite successful at measuring risk attitudes. Its success is that the specific risk parameter it estimates in one institution is consistent with the risk parameter it estimates in the other institution.

The second type of behavior under uncertainty that I am concerned with is the situation where a worker does not know if he will be monitored by his employer. Past economic experiments have shown that some subjects work without financial motivation. This result is in opposition to the traditional economic theory that an individual would not exert costly effort unless he were to be compensated.

I investigate this issue by designing an experiment with random monitoring. The way I look at individual behavior under uncertainty is by having individuals participate in an experiment where the incentive structure is similar to a work environment. The subjects are told that if they complete some task, they will be assured of earnings. The subjects face uncertainty in the experiment because they do not know if they will be monitored. This means that there is some possibility that their work will be checked, but they only know the probability at which this occurs. The way I have designed my experiment allows for subjects to make money outside of the money they earn for completing the specified task. Under one scenario, the incentives allow a worker to get paid whether or not he does his job and he can make himself better off by doing something else (shirking); he has an incentive to not do his job. In past experiments, the way a subject shirked was by sitting still. So, if a subject is bored by sitting still, shirking may actually be more costly to them than working on the task. My design allows me to test if the past experimental results are artifacts of the past experimental methodology.

In chapter four I explain specific features of the institutional environment that I have designed to mimic a work setting with an outside option. I then report the results of subject behavior in this experiment under various incentive schemes. What I find is that subjects do shirk when they are given the incentive to do so. But monitoring reduces shirking. When there is no monitoring, subjects in my experiment make use of the money making opportunity they have outside of completing their main task. But when they are monitored, this behavior decreases. I also find that requiring too much work may be counter-productive. By requiring too much work, people may respond by offering less effort than they would if the requirement was lower.

CHAPTER 2

Literature Review

2.1 Methodology

This dissertation will make extensive use of experimental methodology. In using such a tool, it is important to understand its benefits and limitations. It will be important to understand these steps because key issues that I deal with question if certain experimental results are due to a failure of valid methods or are truly descriptive of individual behavior.

Roth (1987) points out three specific uses for experimental economics. They are: *Speaking to theorists, Searching for facts (meaning),* and *Whispering in the ears of princes.* Speaking to theorists is concerned with the idea of theory testing. We can think of this in terms of prescriptive and descriptive economics, where prescriptive refers to what theory predicts people should do, and descriptive refers to what people actually do (for my purposes, this will mean what they are actually observed to do in the laboratory). Searching for facts applies to situations where there does not yet exist mainstream theory and by examining results from the laboratory, experiments can be thought of as informing theory. This is similar to the Yogi Berra saying: "You can observe a lot just by watching."[1] My first experiment that tests for consistency of risk attitudes will address both of these points. I will consider the general question about the consistency of risk preferences, but I will go beyond that and test if a particular model is successful in deriving consistent risk parameters. Whispering in the ears of princes suggests that experimenters may be able to use their results to inform policy

[1] Taleb (2007) mentions that this quote may be apocryphal and actually belong to Niels Bohr.

decisions.[2] The results from my monitoring experiment will point out that monitoring does work, but that managers will want to pay careful attention to the requirements they are setting for their employees.

Economic experiments allow for control of variables that are always or at least often unknown in a natural setting. Smith (1962) runs an experiment called the double auction where he places subjects into two groups (buyers and sellers) and has buyers announce prices they are willing to pay and sellers announce prices they are willing to accept. What drives these bids (willingness to pay) and asks (willingness to sell) is the underlying values that Smith has induced. This is the key aspect of control. Without this aspect of control, market observations only allow us to see what happens, but not the underlying variables that are causing a certain outcome to happen. Smith's results are noted for their consistency with the supply-and-demand model. Samuelson (2005) suggests that economic experiments were able to gain their initial footing due to showing that laboratory market results are consistent with fundamental economic theory.

Another advantage of using economic experiments is that they are replicable. This allows other researchers to verify and also test variations on elements of a given experiment. By replicating past experiments, one can show that the incentives of the experiment are what is driving the results of the experiment. Once past results are verified, modifications can be made to the past experiment to investigate new questions. The common practice in experimental economics is for experimental instructions and data to be made freely available. In this regard economic experiments merely follow the lead of the traditional sciences. Beveridge (1950) writes that the traditional method (Beveridge here is speaking about natural science experiments) is to have similar groups and to "vary one thing at a time and make a note of all that you do." This is important on two levels when we want to transfer the concepts from natural science experiments to economic experiments. First, it allows for the *ceteris paribus* condition to be met within the framework of the initial experiment. Second, it allows for modification of instructions for future experiments. Unlike natural science experiments, there is sometimes the question of whether or not subjects participating in an economic experiment are being biased by particular instructions. When

[2] Policy is most commonly used to denote government policy, but can more broadly be interpreted. For my purposes, I will develop an experiment that could inform the policy of a firm.

this question arises, the same experiment can be run with modified instructions, and we can see if past results are robust in the face of this change. It should be noted that economic experiments do not permit deception. What the instructions state as true is true. One reason for this is that if subjects know that there is no deception, they know they are responding to the incentives in the instructions and not hidden incentives. A more practical reason is that lying to the subjects could effect their willingness to participate in future experiments, or at the least they would not trust future instructions. This does not mean that subjects have full information, it simply means they are not lied to. For example, in my experiment, when subjects are told that something happens with a 15% probability, it happens with 15% probability.

Besides changing instructions, one could also change the makeup of the subject pool chosen to participate in the experiment. Currently, most research in experimental economics uses university students as subjects. This practice has lead to the critique that this commonality may bias experimental results (*Why should we expect university students to act like professionals?).* If it is the case that results are biased, the remedy is simple, the experimenter can re-run the experiment with a subject pool of professionals. However, Davis and Holt (1993) provide a list of studies that show that performance of professionals is in line with that of university students. One particular example (Smith, Suchanek, and Williams (1988)) finds that both students and professionals are susceptible to bubbles and crashes in laboratory asset markets. Guillen and Veszteg (2006) investigate this issue in terms of demographics by examining the characteristics of over 2000 subjects in 74 different experiments and find that demographic differences account for less than 4% of observed variation.

The subject pool critique is a component of the broader critique concerning the external validity of economic experiments. This critique applies more so when we are dealing with Roth's classification of *whispering in the ears of princes* as opposed to *speaking to theorist* and *searching for facts.* This is because, as Plott (1982) points out, any economic model can be critiqued regarding its external validity. When experiments are used for theory testing they are set up to give the theory its best chance at success. The idea is that if the theory cannot survive a laboratory where it is given its best chance, it will be highly unlikely to survive a natural setting. When experiments are used to search for facts, the external validity critique is not warranted because the experiment is not trying to inform policy, but to inform theory (i.e. experiments are used to help develop theory that can then be scrutinized more closely in the laboratory). So, we are left with questioning why experiments are valid for informing policy. This gets us back to the key benefit of economic experiments, control. Plott (1982) describes scenarios where a prosecutor and regulator have to make decisions about cases to pursue. They may observe data that is consistent with illegal activity, but they are not certain that it is due to illegal activity. By setting up an economic experiment where certain variables can be controlled, a better inference can be made as to whether outcomes that are consistent with illegal activity are actually caused by illegal activity.

A more theoretical defense for external validity is developed in Smith (1982). Smith makes a strong case for what he describes as parallelism: "propositions of behavior of individuals and the performance of institutions that have been tested in laboratory microeconomies apply also to non-laboratory microeconomies where similar ceteris paribus conditions hold." Smith identifies four conditions that lead to parallelism. We can also think of these four conditions as necessary for a valid experiment. They are: nonsatiation, saliency, dominance, and privacy. Nonsatiation means that when all else is equal subject would rather have more of the reward medium than less of it. Saliency means that the money subjects receive in the experiment depends on how they behave in the experiment, and that the reward is high enough to motivate the subjects. Dominance means that potential earnings are greater than any subjective costs of participating in the experiment. Privacy means that a subject only knows his own induced valuations. By ensuring privacy, the experimenter diminishes the possibility that subjects will derive utility from elements other than the reward medium. These last two conditions are necessary for control in the experiment.

The main focus of my dissertation is how people behave when facing uncertainty. One key aspect of this is deciphering an individuals risk preferences. How an individual behaves when facing a particular type of uncertainty allows for his *risk preference* to be classified. He is either risk averse, risk neutral, or risk loving. Knight (1921) distinguishes two types of uncertainty. In the literature, this is referred to as the difference between risk and ambiguity. An individual is dealing with risk when he faces uncertainty, but he knows the odds. An individual is dealing with ambiguity when he faces uncertainty, but does not know the odds. An example of this distinction would be gambling with a fair coin (risk) versus gambling when it is not known if the coin is fair (ambiguity). It is still an open debate as to how much this distinction matters. If individuals form probabilistic beliefs about outcomes and then act, their behavior should be similar in both cases. The experiments in my dissertation will consider instances where the subjects are informed of the probabilities that exists in all institutions (e.g. they know the probability of heads versus the probability of tails). This classification allows for a technical definition of risk and risk preferences.

Imagine that someone is given the choice between 15 dollars for certain, or a gamble using a fair coin. If the coin flip results in heads he gets 10 dollars, if it results in tails he gets 20 dollars. If he is indifferent between this gamble and the certain 15 dollars, he is said to be risk neutral. If he values the certain 15 dollars more than the lottery he is risk averse, and if he values the lottery more he is risk loving. More generally, someone who is risk neutral values things at their expected value. This means that a risk neutral person would have a linear utility function. With a linear utility function, any line that connects two points will be traced directly on top of the utility function. So, if the possible lottery outcomes are the two points under consideration, a straight line connecting them can represent the gamble, and since we have a 50/50 probability between the high and low value, the midpoint of that line would be the expected value. That means, that the utility of the expected value is equal to the expected utility. This is not the case with a concave utility function. A straight line that connects two points on a concave utility function will sit below the curve. This means that the expected utility of the gamble is less than the utility of the expected value, so someone with a concave utility function would prefer the certain $15 over the gamble. Thus, risk aversion is modelled with a concave utility function. The same argument implies that a risk loving person can be modelled with a convex utility function. The straight line connecting two points on a convex utility function would sit above the curve, and the expected utility of the gamble would be greater than the utility of the expected value. So, a risk lover prefers the gamble.[3]

The main issue that I focus on in chapter three is the consistency (or lack of) of individual risk preferences across institutions. I use economic experiments and place subjects in two different institutions to see if they behave similarly in both institutions. I use an auction

[3] For a survey of the origin and development of this concept see Camerer (1995).

institution to derive one estimate, and a lottery to derive the other estimate. I will first discuss particulars of deriving risk preferences from an auction, and then move on to lotteries.

2.2 Background on Risk Aversion in Auctions

The general facts that I discuss about auctions have been compiled from: Kagel (1995), McAfee and McMillan (1987), and Wolfstetter (1995). Simply put, auctions are a rationing device used to allocate goods to the highest bidder (from the standpoint of economic efficiency this would ideally be the person who valued the item the most). There are many types of auction institutions, but I will focus on four particular auctions that will allow me to develop a relationship between auctions and risk preferences. These auctions are: the English auction, the second–price-sealed-bid (SPSB) auction, the Dutch auction, and the first-price-sealed-bid (FPSB) auction.

In the "Japanese" version of the English auction[4], the auctioneer opens the bidding at some price and everyone is in the auction, the auctioneer calls out higher bids and some bidders drop out. This goes on until only one willing bidder remains. The optimal strategy is for a bidder to remain in the auction until his value of the item is surpassed. The winning bidder will pay whatever the second highest bid plus whatever the bid increment is. For example, consider the case where two people remain in the auction. *Bidder 1* (who values the item at \$100) and *Bidder 2* (who values the item at \$120) remain in the auction; assume that the bid increment is five dollars. When the auctioneer calls out a price of \$95 both bidders remain in the auction, at the price of \$100 both would still remain. Once the price reaches \$105 *Bidder 1* will drop out and only *Bidder 2* will remain. So we see that Bidder 2 wins the auction and pays \$105, he is made better of by \$15 (his value minus his winning bid).

In the SPSB auction, bidders place their bids in a sealed envelope and submit them to the auctioneer. Once the auctioneer has all the bids, he opens them and awards the item to the owner of the highest bid. The price the winner pays is the second highest submitted bid.

[4]The more commonly discussed form of the English ascending auction has been shown by Isaac, Salmon and Zillante (2007) to have a more complex optimal bidding strategy. This stems from the bidders in an English auction having the ability to bid above the bid increment (this is referred to as jump-bidding).

The optimal strategy in this auction is for the bidders to submit their true value. We can see this by focusing on the alternatives. If *bidder 1* bids below his value and wins, nothing changes. He pays the same amount he as if he had bid his true value. If he bids below his value and loses there is the potential that he has foregone earnings. If a bid of his value would have been enough to win, by bidding below his value he forgoes his value minus the winning bid. If he bids above his value and wins, he either pays the same price he would have if he had bid his value, or he actually loses money. If the second highest bid is above *bidder 1's* value, then *bidder 1* loses the second highest bid minus his value. So, the SPSB and the English auction have rather straight forward strategies. Also consider that if the bid increment is small enough in the English auction, the price paid by the winner would be the same as in the SPSB (In the English auction the winner pays his bid, but his bid is the second price plus the bid increment) .

In the Dutch auction, the auctioneer starts the bidding by quoting a price higher than he believes any of the bidders would be willing to pay. He then lowers the price until someone is willing to pay. The first person announcing that he is willing to pay wins the auction. The winner pays his bid. The optimal strategy is not as straightforward in this auction. If the bidder claims the item with the simple strategy of bidding his value, he is not made better off (his value minus his bid is zero). In order to make himself better off he has to win the item at a price below his value. The longer he waits to claim the item, the larger his potential profit (value minus bid). However, it is also the case that the longer he waits, the more likely it is that a rival bidder will claim the item. This is the risky choice in an auction. Someone who is risk averse is going to want to protect the profit more than someone who is risk neutral. So, they will not wait as long as a risk neutral bidder to claim the item. The optimal strategy in this auction is for the bidder to bid some amount below his value. This will be stated formally later in this section.

The protocol in the first-price-sealed-bid (FPSB) auction is similar to the SPSB, but the winner pays his own bid, not the second highest bid. The optimal strategy for the FPSB is similar to that of the Dutch auction. A bidder will only make a profit if he bids below his value, but by doing so he lowers the likelihood of winning the auction. We can now move on to a more technical discussion of what an optimal bid strategy will be in this scenario. We will see that it is possible to get an estimate of how risk averse a subject is by how much they hedge their bid in a FPSB auction. It will then be possible to compare this estimate with the estimate that I will elicit from another institution.

Vickrey (1961) solved for the optimal bid strategy given certain assumptions. His model assumes that all bidders are risk neutral and that the distribution of their value space is known. A bidder already knows his own value, but he only knows the range of the values that his rivals could have (and that these values come from a uniform distribution). It is also the case that a bidder knows how many rivals he faces. Given these assumptions Vickrey shows that the optimal bid is for the bidder to bid some constant proportion of his value dependent on the number of bidders in the auction. As the number of bidders increases, optimal bids also increase.

Vickrey, Myerson (1981), Riley and Samuelson (1981) show that the above four auctions all yield the same revenue (given the assumptions listed above). This is referred to as the Revenue Equivalence Theorem. Holt (1980) shows that in a FPSB auction, if someone is risk averse, they will bid higher than the RNNE.[5] Cox, Roberson and Smith (1982) show, in a laboratory setting, that bids in a FPSB auction are generally above the risk neutral prediction. This would mean that revenue equivalence would not hold (FPSB auctions would generate higher revenues than the English auction).[6] In the same year, Milgrom and Weber (1982), with an argument similar to Holt (1980) show generally that individual who are risk averse will bid above the risk neutral bid. Cox, Roberson and Smith develop a model to deal with this observed overbidding, but it appears that this was a concurrent finding (as neither Milgrom and Weber nor Cox, Roberson and Smith cite each other). Cox, Roberson and Smith state that they are building their model to consider risk aversion on the suggestion of John Ledyard. The model they develop has come to be known as the constant relative risk aversion model (CRRAM). CRRAM allows for the risk preferences of the bidders to be backed out from their observed bidding behavior. Cox, Smith and Walker (1988) further explored risk preferences in FPSB auctions using the CRRAM. This model will be discussed

[5] Holt was studying procurement auctions which have the property that the lowest bidder wins the auction (e.g. contractors bidding to provide a service). He shows that risk averse bidders would bid less than risk neutral bidders. This means that in an ascending auction (as opposed to a procurement auction), risk averse bidders will bid above the RNNE.

[6] Ivanova-Stenzel and Salmon (2006) show that when entry decisions (i.e. subject can choose to enter or not, so the number of bidders is not fixed) are considered the FPSB auction may not generate significantly higher revenues.

in detail in chapter 3.

Harrison (1989) questioned whether or not the experimental methodology used by Cox, Smith and Walker met the requirements for a valid experiment that were laid out in Smith (1982). This led to a debate that John D. Hey (1991) said "stirred the passions of the experimental community." Harrison posited that the deviations from risk neutral behavior that Cox, Smith and Walker observed were too small in the payoff space to have any significant meaning. This is called the flat maximum critique. If we think of subjects responding to the incentives of an experiment to make money, we can imagine that they have a function that they are trying to maximize. The flat maximum critique is an issue when subjects can reach the maximum with a range of behavior as opposed to a distinct behavior. Cox, Smith and Walker (1988) do not have a flat maximum, but Harrison argues that the payoff space differences are so small that the incentives in their experiment are not salient. His point was that there was a loss of dominance since deviations from the RNNE were not costing subjects much money. Cox, Smith and Walker (1992) and Friedman (1992) point out that there are technical shortcomings with Harrison's argument. Cox, Smith and Walker argue that Harrison attaches cardinal value to ordinal utility, and that Harrison does not get to be the arbiter of what is a "small difference" in the payoff space. Furthermore, Friedman points out that if what Harrison says is true, then deviations from the RNNE should fall on both sides of the RNNE, but Cox, Smith and Walker (1988) report that subjects go in one direction (they consistently bid above the RNNE).

There are alternative explanations for overbidding in a FPSB. These critiques are different than Harrison's. To simplify, Harrison was saying that overbidding really wasn't overbidding, just faulty design. An alternative to this is to acknowledge overbidding, but argue the causality. That is to say the methodology is fine, but the interpretation and model may not be right. The first to make this argument were the authors of CRRAM themselves when they suggested that subjects possibly receive joy from winning the auction over and above the monetary gain. They later show that this model (Joy of Winning) does not do as good a job of explaining behavior as the CRRAM (Cox, Smith and Walker (1988)). Engelbrecht-Wiggans (1989) studies the effect regret may have on bidding behavior. Subjects can have two types of regret. If a subject wins the auction, he may feel that he has paid too much and could have still won by bidding less. If a subject loses the auction, it is possible that he could have won if he had bid higher (but still below his value). If this second effect dominates, overbidding could be the result. Friedman (1992) considers the possibility that subjects just start out with some ad-hoc rule (bid some percentage of their value) and if given enough time would eventually correct to the Vickrey RNNE. These arguments share the commonality of accepting that overbidding occurs, but find the risk aversion claim dubious. Cason (1993) and Kagel and Levin (1993) each use different institutions and do not find the majority of their subjects to be risk averse. It should be noted that the institutions used by Cason, and Kagel and Roth are more complex than the FPSB auction. Kagel (1995) claims that the complexity should not matter when trying to gauge behavior. I understand his argument to be something like: people do not need to know economic theory to conform to it, and if they don't conform to it, then something is amiss with the theory. The problem with this argument is that complexity can take two forms. In one form Kagel's point is valid, but in the second it may not be. A task can be complex in that it is difficult for subjects to figure out what to do even if they understand all the rules, or a task may be complex because the subjects cannot figure out the rules. This latter case, where subjects may not be sure of the given incentives could lead to questionable results. We will see below that Kagel (2001) makes a similar argument in a different context. He will clarify the incentives in an experiment and in turn get results that are at odds with past experiments.

In this section, we have seen the proposition that overbidding in FPSB auctions is due to risk aversion. There have been critiques that overbidding doesn't really mean anything (Harrison (1989)), overbidding does mean something, but it isn't necessarily caused by risk aversion (Friedman (1992) and Engelbrecht-Wiggans (1989), Cason (1993) and Kagel and Levin (1993)). In the next section I will further explore the risk aversion aspects in the FPSB auction and introduce institutions that have been used to elicit subjects' risk preferences.

2.3 Procedures for Estimating Risk Aversion

One way to elicit someone's risk parameter is to ask him how much he would be willing to pay for a lottery. We have seen that someone who is willing to pay less than the expected value is risk averse. The problem with this in a market setting is that it requires truthfulness on the part of the buyer. If you just ask a person how much they are willing to pay for something, he may submit a lowball offer. One way to get around this is to ask hypothetically. Now there is no reason to lie, but there is also no reason to be truthful (we have lost the salient reward precept). In an attempt to get around these issues, Becker, Degroot, Marschak (1964) devised what is referred to as the BDM mechanism.

The BDM mechanism is similar to the SPSB auction discussed above, but there are some key differences. In the BDM the subject is endowed with a lottery, and is then asked for the minimum price that he would be willing to sell the lottery back to the experimenter. The experimenter then draws a random number. If the number drawn is greater than the stated min selling price, the lottery is sold, and the subject receives the value of the random number. If the random number is lower than the min selling price, the subject plays the lottery and earns the proceeds. This is similar to the SPSB auction because the optimal strategy is for the stated min price to be the subject's value. If subjects value their endowed lottery by more than its expected value, they are said to be risk loving.

Harrison (1990) uses the BDM mechanism to back out his subjects' risk parameters with the CRRAM. He uses these results to further question the results found by Cox, Smith and Walker (1988). He shows that the risk parameters backed out from a different institution (BDM) yielded vastly different estimates. He found that his subjects were much more risk loving. Harrison's argument is that if subjects are risk neutral/loving in the BDM (as he found them to be) and preferences are stable (which is commonly assumed) then Cox, Roberson and Smith's model is not valid. However, an alternative exists. It could be the case that the different results were due to Harrison using a different subject pool, or that preferences are not stable. Isaac and James (2000) explore this issue by having the same group of subjects participate in both institutions. On aggregate, they find that their subjects act similarly, in the first price auction, to the subjects in Cox, Smith and Walker, and that their subjects act similarly, in the BDM, to the subjects in Harrison (1990). Isaac and James calculate the risk parameter for each individual subject in both institutions, further they rank the subjects in each institution (from most to least risk averse). Their results show that neither values nor ranks are preserved across institutions. This raises a few questions. Do subjects truly have unstable parameters, or is something else at play? Do either of these devices measure risk aversion? Is the technique used to back out the risk parameter valid, is it valid for one institution and not the other?

The model used by Cox, Smith and Walker and followed by Harrison (1990) and Isaac and James assumes that individuals have constant relative risk aversion (CRRA). Kagel et al. (1987) and Smith and Walker (1993), in contradiction to this, show that as the expected payoff of winning the auction increases subjects' bids increase by more than the CRRAM would predict. Subject behavior in these studies is consistent with risk aversion, but it is increasing relative risk aversion not CRRA. Goeree, Holt and Palfrey (2000) explore alternatives explanations of overbidding by trying to estimate parameters with various models. They find that the risk aversion model is the best fit.

There is strong evidence that subjects in first price auctions behave, at the least, 'as if' risk averse (we saw earlier that it is not fully agreed upon that risk is driving behavior in the FPSB auction but that subjects overbidding is at least consistent with risk averse behavior), and it is possible to measure this behavior without assuming either constant relative or increasing relative risk aversion.

Holt and Laury (2002) introduce a new lottery device that also allows for elicitation of risk preferences. It shares similarities with the BDM discussed above, but there are some important differences. In the BDM subjects are specifically asked how much they value a gamble and are then incentivized to be truthful. In Holt and Laury (HL) subjects are presented with a series of two lotteries and asked to choose the one that they prefer. This gets around any possible confusion a subject may have in calculating a value himself. It is not a question of how much, just simply pick one or the other. By observing their decisions over a menu of lottery pairs, it is possible to gauge their risk preference.

Similar to Kagel et al (1987). and Smith and Walker (1993), Holt and Laury find that subjects' risk aversion increases as potential earnings increase. A nice feature of the Holt-Laury lottery is that while it can be used to get an estimate of subjects risk parameters under various assumptions of subjects' utility functions, it also allows for some proxy of risk aversion without assuming an underlying parametric form.

2.4 Background on Principal-Agent Model and Experiments

When designing a contract it is often the case that an information asymmetry exists. In the principal-agent problem, this occurs when an employee (the agent) can hide information from his employer (the principal). Technically, this is referred to as hidden action or moral hazard. The problem arises because there is some uncertainty between the agents effort and the outcomes that the principal observes. The importance of this is that the principal cannot be certain that the agent is shirking (not working as hard as he could) by simply observing outcomes. A hard worker could have a bad outcome due to bad luck, or a lazy worker could have a good outcome due to good luck. There is a long standing and mature theoretical literature dealing with the principal-agent problem (Mirrlees (1975), Grossman and Hart (1983)). In the basic model, the principal has to offer the agent a contract that will motivate his participation and effort. If both the principal and agent are risk neutral; the solution in the case where the principal does want the agent to exert high effort is a 100% commission contract with a fixed payment from the agent to the principal. We can think of a wholesaler selling to a retailer as opposed to taking a percentage once the item sells to the end user. The solution changes when the agent is risk averse. A risk averse agent is not willing to pay in full for the product because he does not like the uncertainty that he faces when he tries to resell the item. In this case, the principal has to pay some flat wage to bear some of the risk (this can be thought of as insurance for the agent. He knows he will get something for his efforts even if he does not make the sale).

An alternative to the principal insuring the agent is for the principal to monitor the agent. If it costs less for the principal to monitor the agent than it does for him to insure his wage, monitoring will be preferred. It will also be important to consider the cost of monitoring versus the amount of money being lost to employee shirking. Recent empirical work has tried to capture how much firms may be losing due to their employees being rational cheaters. Nagin (2002) defines a rational cheater as someone who will shirk when the marginal benefit of doing so exceeds the marginal cost. A rational cheater will exert low effort on the job if he thinks he can get away with it. Some reports estimate that shirking workers cost employers billions of dollars in productivity losses yearly[7]. Employers (principals) who are aware of the financial incentives they are giving employees (agents) may introduce a monitoring system with performance goals to alleviate the perceived problem. These issues with worker motivation are difficult and complex. Consider a recent New York City court case:

> On March 9, 2006, John B. Sooner, a New York City administrative law judge, recommended that Toquir Choudhri, a 14-year veteran of the city Department of Education, receive only a reprimand for disobedience, even though supervisors wanted him fired for using the Internet for personal matters[8]. Spooner wrote that Choudhri credibly stated that he completed all assignments given to him by his boss and used the internet while he awaited further assignments. These statements were corroborated by the absence of proof that Choudhri was ever criticized for poor productivity or for not completing specific assignments.[9] The New York City Chancellor of Education, Joel Klein, decided to fire Choudhri anyway. Klein stated that 'the penalty of termination is appropriate and not shocking to one's sense of fairness, ... Choudhri's abuse of the Internet at the time he is supposed to be performing his job demonstrates his disinterest in the job.'[10]

The worker in the above case was fired for shirking on the job when his employer found him surfing the internet. The worker did not think he deserved to be fired because he had completed all of his assignments. The worker thought he was being monitored in regard to fulfillment of some quota, and he had fulfilled his quota, but the employer disagreed. This case demonstrates the problems caused when the monitoring system is not well delineated,

---

[7] Frauenheim, Ed. "Stop Reading This Headline and Get Back to Work." *CNET,* Monday, July 11, 2005, http://news.com.com/Stop+reading+this+headline+and+get+back+to+work/2100-1022_3-5783552.html

[8] Klopott, Freeman. "Should You Be Fired for Using the Internet While at Work?" *PC World,* Tuesday, May 02, 2006, http://pcworld.com/article/id,125597-page,1/article.html?RSS=RSS

[9] Department of Education v. Choudhri, New York City Office of Administrative Trials and Hearings, No. 722/06 (3/9/06)

[10] Associated Press. "NYC Fires Man For Web Surfing At Work." *CBS News,* http://www.cbsnews.com/stories/2006/05/06/tech/main1596034.shtml

but it also shows how concerned some employers are about any behavior consistent with shirking.

Past experimental research (Cadsby et al. (2004), Dickinson and Villeval (2005)) has shown that some workers in a laboratory setting work without incentives. This shares some similarities with Akerlof's (1982) proposition of a gift-exchange where an employee may choose to exert high effort because he is grateful for the "high" wage that the principal is paying him. This phenomenon has been broadly studied in terms of its reciprocity and trust aspects (Fehr et al. (1993), and Berg et al. (1995)). With a more narrow focus, Fehr et al. (1998) use experiments to model the gift-exchange in terms of a "fair-wage," and find support for the idea that high wages can lead to efficiency gains and better outcomes for both the employer and employee. There are some critics to this work. I can best frame the criticisms in terms of how the past gift exchange experiments do not follow the precepts for a valid experiment. Taken on their face value, the gift-exchange experiments reject traditional agency theory. This falls under the framework of using experiments to speak to theorists. But Engelmann and Ortmann (2002) show that the earlier work did not meet the necessary requirement of giving the theory its best chance for success. Recall the underlying thinking, theories that fail in the laboratory when given their best chance for success are highly unlikely to be validated in a natural setting. It is not the case that theories given a poor chance at success can be invalidated. Engelmann and Ortmann change some features of the past experiments and find that traditional theory performs well. Charness, Frechette and Kagel (2001) show that the past experimental failure of agency theory captured by the gift-exchange experiments could be caused by something as simple as the subjects not realizing the incentives of the game. By providing their subjects with a comprehensive payoff table, the gift-exchange finding is severely lowered and traditional theory performs well. So, even when the theory is not given its best shot in terms of design, simply clarifying the instructions is enough for traditional theory to survive.

It is important to investigate if Cadsby et al. (2004), and Dickinson and Villeval (2005) laboratory results indicate behaviors we would see in a real work setting, or if the possibility exists that these observations are an artifact of the experimental design. Cadsby et al. (2004), and Dickinson and Villeval (2005) were not specifically looking at effort with low incentives, so there is remaining value to their research even if the observations of subjects working without incentives are an artifact of their design. Below I will further discuss the aspects of their work that show effort when their subjects are not motivated. But I note here that Cadsby et al. are interested in how people self select into various payment schemes; Dickinson and Villeval are concerned with testing competing theories related to the principal-agent problem. They test if monitoring leads to increased effort or if announcing a monitor system informs subjects that they are expected to be shirking. Even if there results are biased towards high effort, they can still gain insight into their primary question.

A common approach used in the laboratory to investigate the principal-agent problem is to give subjects a cost function and have them choose some effort level. Nalbantian and Schotter (1997), for example presents an experiment in which a subject is monitored with probability. If the subject is not expending a certain level of 'effort', he will be terminated. The 'effort' in this case is not physical exertion but rather a figurative effort. This number they pick is indeed interpreted as effort and therefore has the property that effort is now explicit. While this matches clearly with their models, it is not clear that subjects perceive this choice as analogous to physical or mental exertion. The (e) chosen by the subject is costly to the subject but it is possible that this is too abstract to model real work.[11] Putting the workers through a real effort experiment will allow an answer as to whether simply choosing effort garners the same behavior as exerting effort, and if it does not, one can be confident that the real effort experiment is a better proxy for the workplace. The approach of using real effort tasks to investigate workplace behavior is becoming more popular (see: Dickinson, (1999); Sillamaa, (1999); van Dijk, Sonnemans and van Winden, (2001); Gneezy, Niederle and Rustichini, (2003); Dickenson and Villeval (2005); Montmarquette, Rulliere, Villeval and Zeiliger, (2004); Falk and Ichino, (2006)).

While an agent's outside option can be represented rather easily in a theoretical framework, it is not so trivial to do in a laboratory. To think of this one must place himself in the position of the subject in the experiment. The subject arrives at the laboratory and is assigned a computer terminal. He is given the option of engaging some task and earning X or he can take the outside option (sit still) and earn Y. If the disparity between the outside option and the participation option is not large, there exists the possibility that the subject

[11] Some pros of figurative effort include control of subject ability, and ease of implementation in the laboratory.

will engage the game to avoid boredom.

The aforementioned research by Cadsby et al. (2004), and Dickinson and Villeval (2005) found that subjects contribute effort even when they have no financial reason to do so. The idea of a moral imperative not to shirk is given as a reason for this behavior by Dickinson. Cadsby's experiment gives subjects seven scrambled letters and the subjects are instructed to make as many words as possible in a given time period. The subjects are allowed to choose a piece rate or a flat rate scheme, the piece rate will pay them per word, and the flat rate will give them some stated amount with no requirement on word creation. These incentives should lead subjects who think they are endowed with word creation ability and low effort cost to choose the piece rate scheme and the subjects who have high effort cost or are not good at word creation to choose the flat rate scheme. The interesting observation is that there are significantly more than zero words created by the subjects who choose the flat rate scheme. This is contrary to the incentive structure, and one would not believe the subjects feel a moral imperative to unscramble letters. Dickinson and Villeval (2005) use a real effort task and they observe that some of the subjects (25%) contribute at or above the desired output level even when monitoring is set to zero. Dickinson, as noted earlier, suggests this as either intrinsic motivation or integrity and commitment to moral principles. It is possible that subjects feel a moral imperative because they are interacting with a human principal, but that is not the case in Cadsby (2004). The intrinsic motivation argument is plausible in both experiments. The subjects might enjoy unscrambling letters, and they might enjoy moving along a curve to get a high value, but it is not clear that observations made under these conditions should be interpreted as analogous to intrinsic motivation one might experience in the real world. In both experiments, the subjects could engage in effort, or do nothing. It is possible that they were engaging the tasks because they were bored. This is similar to Choudri's claim that he browsed the internet only because he had no other work to do. If the subjects in the experiment do not engage the task, they have nothing else to do.

Other experiments have shown that outside options have important effects. Lei, Plott, and Noussair (2001) show that excess trading in an asset market can be reduced by giving subjects something to do beyond trading in the asset market. Pevnitskaya and Palfrey (2004) show that over-entry into an auction can be reduced by allowing subjects an outside option of a computerized version of rock-paper-scissors. Van Dijk et al. (2001) conduct an experiment where subjects are enabled to work on two tasks in the same period. The earnings from one task go in to a group account while the earnings from the other task go in to a private account, similar to a public goods game. This alleviates the effort only due to boredom problem that is plausible in the Cadsby(2004) and Dickinson (2004) papers. The specific task the paper uses has the subjects search a grid looking for the highest payoff. The idea of having two of these for the subjects to play cures the boredom critique.

2.5 Going Forward

I began this literature review by describing reasons for using experimental economics and methods for conducting a valid economic experiment. The following two chapters will use the methods outlined and show results that *speak to theorists*, *search for facts*, and could be used to *whisper in the ears of princes*.

For my experiments, the reward medium is money. All subjects are paid a $10 show up fee, but this is not enough to ensure that they are sufficiently motivated. To meet the saliency requirement, the subjects' actions in the experiment must be tied to the reward medium. So, subjects have a salient reward to attend the experiment, but they also have salient rewards in the experiment.

In Chapter 3, I will report the results of an experiment where I have the same subjects participate in two institutions. I use the FPSB auction and the Holt-Laury lottery discussed above. I do this to see if subjects have consistent risk preferences between these institutions. By using these specific institutions, I can evaluate risk preferences without assuming an underlying utility function. I am also able to evaluate how well the CRRA model does in measuring risk preferences across these institutions. So, the first result is a search for facts, and the second is speaking to theory.

In Chapter 4, I examine individual behavior in an environment designed to mimic work for an employer. This chapter develops a theory that is a derivative of the basic principal agent model, specific to the situation that I am modelling. I then test this theory in an economic experiment. The success of the theory test gives insight into how managers may want to motivate their employees. So, this chapter speaks to theorists and could be used to inform policy. Further, it is my intention that this chapter also speaks to experimentalists by addressing a methodological issue. I examine if specific features of past experimental designs were responsible for influencing behavior in a way that past experiments did not take into account.

2.6 Conclusion

I have used the experimental economics methodology laid out by Smith (1982) to investigate how individuals respond to uncertainty in various settings. Recall that Roth pointed out three specific uses for experimental economics. I will summarize my results by discussing how they fit into Roth's framework, and give some final remarks.

First I will discuss how my dissertation *speaks to theorists,* and *searches for facts.* I am combining these two uses because my risk preferences experiment addresses both. The basic question regarding consistency of risk preferences can be thought of as a *search for facts,* but it is heavily intertwined with economic theory. It was not my intention that my results would make a strong case for economic theory, but it has turned out that way. In examining risk preferences across institutions I have found that individuals behave quite consistently. I was able to use two devices that allow for non parametric results to show that individuals who are observed to be risk averse in a first-price-sealed-bid (FPSB) auction are likely to be risk averse in a Holt-Laury (HL) lottery. I consider this as a *factual* result that can inform theory. This is the concept Roth described when speaking about a *search for facts.* This is a nice result for economic theory because it allows for the individuals to be modelled more generally than if their characteristics changed due to the institution. I have not however closed the question that was raised by Isaac and James (2000) that subjects bid 'as if' risk neutral/risk loving in the BDM device. But the fact that my results are supportive of traditional economic theory suggests that focus should be placed on examining what is going on in the BDM. As for *speaking to a specific theory*, my results were supportive of the CRRA model that was put forth by Cox, Roberson and Smith (1982). I found a strong relationship between risk parameters backed out of the FPSB auction and those backed out in the HL lottery.

In my monitoring experiment, I found that subjects respond similarly to what my model predicts, and since my model is derivative of basic principal-agent theory for my specific institution, economic theory again performs quite well.

My monitoring experiment also *whispers in the ears of princes,* if one is bold enough to equate employers with princes. The main point of this concept is that economic experiments can be used to inform policy. My experiment suggests that managers should, in fact, monitor their employees if they are giving them incentives that induce shirking. However, my results would inform managers that they should be wary of requiring full out effort all of the time, as my results showed that subjects became much more likely to shirk when their task became increasingly taxing. The danger with this for managers is that by requiring too much, they may actually get less than if they required a more modest output.

The last result that I would like to discuss could be termed *speaking to experimentalists.* My monitoring experiment shows that when individuals are given an incentive to shirk, they do. This is consistent with economic theory, but at odds with some past real effort experimental results (Cadsby(2004) and Dickinson and Villeval (2005)). Since the past papers were not focusing on this issue, some briefly claimed that this could be due to subjects maximizing a utility function that has parameters theorists may not generally model. This could, in fact, be true, but I would claim that it should not be the first response. I believe the first response should be a self critical examination of the experimental design. It is my contention that what was dubbed as a moral imperative to not shirk could simply be a loss of control in the experiment. If subjects enjoy participating in the experiment as opposed to sitting still, the experimental design should take pains to ensure that 'costly' participation overrides the cost of sitting still.

CHAPTER 3

First Price Auctions, Lotteries, and Risk Preferences Across Institutions

3.1 Introduction

Recent experimental research has called into question the assumption that individuals who are risk averse in one institution will be risk averse in another institution. If these findings are robust, and individuals' risk preferences in one situation have no bearing on how they act in other situations, economists will face a difficult task in trying to model behavior where individuals face uncertainty. If, however, there are possibly classes of institutions, or specific features of institutions that lead to predictable behavior, the situation is much less dire.

Much of the experimental economics risk preference analysis can be traced back to Cox, Roberson and Smith (1982), and Cox, Smith and Walker (CSW) (1988). In studying first price auctions, subjects were commonly observed to bid above the risk neutral Nash equilibrium prediction. A plausible explanation is that the subjects are risk averse (Milgrom and Weber (1982)). By bidding higher, a subject increases the probability that he will win the auction, albeit at a higher price. CSW further showed that they could back out a subject's risk parameter given his bid.

Harrison (1990) put forth a challenge to the risk parameters derived by CSW. Using the BDM mechanism (Becker, Degroot, Marschak (1964))[1] and the same model as CSW, he

[1] The BDM works as follows: a subject is endowed with a lottery, and asked his min price of selling this lottery back to the experimenter. The experimenter then draws a random number. If the number drawn is greater than the selling price, the lottery is sold and the subject receives the value of the random number. If the random number is lower than the min selling price, the subject plays the lottery and earns the proceeds.

found that his subjects' behavior was more risk loving than the behavior of CSW's subjects. It is possible that the risk parameter derived in Harrison (1990) is different than that derived by CSW simply due to having different subjects. Isaac and James (2000) explore this by having the same subjects participate in both institutions. On aggregate, they find that their subjects act similarly, in the first price auction, to the subjects in CSW, and that their subjects act similarly, in the BDM, to the subjects in Harrison (1990). Isaac and James (2000) calculate the risk parameter for each individual subject in both institutions, further they rank the subjects in each institution (from most to least risk averse). Their results show that neither values nor ranks are preserved across institutions. This raises a few questions. Do subjects truly have unstable parameters, or is something else at play? Do either of these devices measure risk aversion? Is the technique used to back out the risk parameter valid, is it valid for one institution and not the other?

The model used by CSW and followed by Harrison (1990) and Isaac and James (2000) assumes that individuals have constant relative risk aversion (CRRA). Kagel et al. (1987) and Smith and Walker (1993), in contradiction to this, show that as the expected payoff of winning the auction increases subjects' bids increase by more than the CRRA model would predict. Subject behavior in these studies is consistent with risk aversion, but it is increasing relative risk aversion not CRRA.[2] Goeree, Holt and Palfrey (2000) explore alternatives explanations of overbidding by trying to estimate parameters with various models. They find that the risk aversion model is the best fit.

There is strong evidence that subjects in first price auctions behave, at the least, 'as if' risk averse, and it is possible to measure this behavior without assuming either constant relative or increasing relative risk aversion. Holt and Laury (2002) introduce a new lottery device that also allows for elicitation of risk preferences. Similar to Kagel et al. and Smith and Walker, they find that subjects risk aversion increases as potential earnings increase. A nice feature of the Holt-Laury lottery is that while it can be used to get an estimate of subjects risk parameters under various assumptions of subjects' utility functions, it also allows for some proxy of risk aversion without assuming an underlying model. In my experiment, I have the same subjects participate in the first price auction and a the Holt-Laury lottery. By

[2]For a detailed account of this debate see Kagel (1995).

using these institutions I can evaluate risk preferences without assuming an underlying utility function. I am also able to evaluate how well the CRRA does in measuring risk preferences across these institutions. I find that subjects behave consistently across institutions. I also find that the CRRA estimates in the first price auction are correlated with CRRA estimates derived from the lottery.

3.2 Experimental Design

The experiment conducted for this paper was designed to elicit subjects' risk parameters in two institutions. The first institution the subjects participate in is a Holt-Laury lottery using the protocol established by Prasad and Salmon (2007). The second institution that the subjects participate in is an independent private values, first-price, sealed bid auction. All subjects participated in the lottery prior to participating in the auction. Economists are often concerned about order effects in experiments. The main concern with ordering is that participation in the first experiment will influence behavior in the subsequent experiment. This is a valid concern when the institutions are closely related (e.g. the ultimatum game and the dictator game, first price auctions and second price auctions) but that is not an issue here, the experiments in this paper do not share common elements. There is also a concern about wealth effects (how money earned in one experiment may affect behavior in another). That is also not warranted for this study because money earned in the lottery does not carry over into the auction.

3.2.1 First Price Auction

In this portion of the experiment the subjects participate in 30 rounds of first price auctions. Each auction lasts 30 seconds, and there are 30 seconds between each auction. The subjects were randomly placed into groups of four. The subjects do not know who the other members of their group are, but they do know that the group consists of the same four members for all 30 rounds.

In each round the subjects draw a number from the uniform distribution [0, 100]. This number is their value for the hypothetical object being auctioned off. The subjects are informed that their value along with all members' values are drawn from a uniform distribution. The subjects are told in the experiment instruction that "it is highly likely that all the members in your group all have different values."

To earn money in this portion of the experiment, a subject must win the auction. The subject who wins the auction receives the difference between his value and the amount he bid for the item. The winner of the auction is revealed at the end of each of the 30 auctions.

3.2.2 Holt-Laury Lottery

In the Holt-Laury lottery institution, a subject is given a menu of lottery pairs (Table 3.1). The subject is to choose either *Option A* or *Option B* for each pair of lotteries. The fourth column in Table 3.1 shows the expected payoff difference of choosing *Option A* over *Option B.*[3] A risk neutral subject will choose *Option A* until decision five where the expected payoff of selecting *Option A* becomes negative, and from this point on would only select *Option B*. A risk loving subject would switch earlier, and a risk averse subject would delay switching until some time after decision 5.

Prasad and Salmon computerizes the Holt-Laury device using z-Tree (Fischbacher(1999)). With this protocol subjects saw each choice individually in a sequential order. The subjects, under this protocol, were not allowed to change their decisions once they had been made. Salmon and Prasad claim that this is not a problem for their purposes as it allows for some gauge of the subjects' sophistication. They later use this gauge to analyze behavior in another situation. The claim is that, with their method, it is not as transparent to a subject that he should only have one switching point. The argument is that sophisticated subjects will be able to pick up the underlying incentives of the lottery and only make one switch. Subjects who make many switches may not be paying close attention to the details of the experiment, and knowing this may allow insight into erratic behavior elsewhere. In the event that subjects do have more than one switching point, the decision where they choose *Option B* and never subsequently choose *Option A* will be viewed as indicative of their risk preference and in this paper will be referred to as a subject's *last switch.* After the subjects

[3]The fourth column (difference in expected value) is not shown to the subjects.

make all ten decisions, one of the ten chosen lotteries is randomly selected and played for actual earnings.

The lottery and auction experiments were conducted over four sessions and 32 subjects participated. The subjects for the experiment were undergraduates at Florida State University. The experiment was computer based and conducted with z-Tree software. When the subjects arrived they were assigned to computer terminals. The instructions for the experiment (see Appendix) were read aloud and the subjects had a chance to ask questions. The subjects were paid a $10.00 show up fee, and were able to earn money based on their performance in the experiment. The average earnings per subject for the experiment were $25.00.

3.3 Theory and Hypotheses

3.3.1 Deriving Risk Parameters Using a First Price Auction.

By assuming that bidders in an independent private values, first-price, sealed bid auction are risk neutral and that their values (v) are randomly drawn from the uniform distribution $[\underline{v}, \overline{v}]$, Vickrey (1961) showed that the optimal bid function $b_i(v_i)$ for a risk neutral bidder in an auction with n bidders is:

$$b_i(v_i) = \underline{v} + \frac{n-1}{n}(v_i - \underline{v}) \tag{3.1}$$

Which can be estimated as:

$$b_i(v_i) = \alpha_i + \beta v_i + \varepsilon_i \tag{3.2}$$

Milgrom and Weber show that individuals who are risk averse will bid above the optimal bid of a risk neutral bidder. The idea is that people participate in an auction to increase their well being, and they do so by gaining some object at a price lower than the value they place on that object. So, their surplus is their value minus their bid (if they win). Low bids increase this surplus, but they lower the likelihood of winning. whereas bids closer to the value of the object decrease the surplus but increase the likelihood of winning. A person who is risk averse will bid higher than a risk neutral person because the risk lies in not capturing the surplus. It follows that the most risk averse person will bid closest to their value, and the least risk averse will bid well below their value. Thus, the estimate of β in Equation 3.2 will be correlated with risk aversion (subjects who are more risk averse will have higher βs. This gives me a non-parametric measure of risk aversion. Cox Roberson and Smith, and Cox, Smith and Walker construct a parametric measure for risk aversion. They observed that subjects do, in fact, typically bid above b_i. They subsequently developed what is known as the constant relative risk aversion model (CRRAM) for first price, private value, single unit auctions. The CRRAM assumes that bidders have heterogeneous risk preferences and that all bidders are aware of this. They model this set of beliefs with equation 3.4 which is similar to equation 3.1 with a slight modification to include a risk parameter r_i. We can see that a more risk averse person (r closer to 0) will bid higher than someone who is less risk averse. Cox, Roberson and Smith show the best response to this belief is to respond with the same bid function so long as $b_i < \bar{b}$. Beyond this point they cannot solve for a closed form linear solution. To see why, we can consider a simple example. Assume you are a bidder in a two bidder auction, you are risk averse (with $r = 0.5$), and assume that you believe your rival to be risk neutral ($r = 1$). If you value the object such that $v = 100$ (let $\overline{v} = 100,$ and $\underline{v} = 0$), and bid according to equation 3.4, you would bid 66.67. However, this would not be a best response. You know that the most your risk neutral rival would bid is 50. If you believed that your rival was slightly risk averse ($r = 0.8$), then the most he would bid is 55.56. In either of these cases, you would want to bid less than 66.67. So, you do not best respond to them by choosing b_i where $b_i > \bar{b}$. In order to apply this to experiment data Cox, Smith and Walker introduce a method, followed by Isaac and James, to censor their data by eliminating observations where the induced value is above $\left(\frac{n-1}{n}\overline{v}\right)$.[4] This forces $b_i < \bar{b}$. The optimal bid function with the CRRAM assumptions in equation 3.4 also allows for the linear estimation shown in equation 3.2.

$$\bar{b} = \underline{v} + \frac{n-1}{n}\left(\overline{v} - \underline{v}\right) \tag{3.3}$$

[4]For my experiment, I will truncate the data by eliminating induced values greater than 75. This is because I have four subjects and the highest induced value is 100 $\left(\frac{n-1}{n}\overline{v} = 75\right)$.

$$b_i\left(v_i\right)=\underline{v}+\frac{n-1}{n-1+r_i}\left(v_i-\underline{v}\right) \tag{3.4}$$

If bidders bid according to the Nash equilibrium, then $\alpha=0$ and we combine equations 3.4 and 3.2, we find β:

$$\frac{\underline{v}+\frac{n-1}{n-1+r_i}\left(v_i-\underline{v}\right)}{v_i}=\beta \tag{3.5}$$

Using $\underline{v}=0$, as I will for the experiment, equation 3.5 allows us to solve for r_i.

$$r_i=\frac{\left(1-\beta_i\right)(n-1)}{\beta_i} \tag{3.6}$$

3.3.2 Deriving Risk Parameters Using the H-L Mechanism

I will use two measures of risk aversion. As noted in Section 2.2, where a subject makes his *last switch* among the ten decisions in Table 3.1 is indicative of his risk preference. Subjects who have a high *last switch* are more risk averse than subjects who have a low *last switch*. I will use this method because it does not require any underlying assumptions to be made about the subjects' utility function. As a secondary measure, I will derive a range of risk parameters for individual subjects using the CRRA model. This will allow for direct comparison of the CRRAM result in the auction. If we assume that subjects possess constant relative risk aversion utility, their preferences can be represented by the following functional form[5]:

$$U(x)=x^r \tag{3.7}$$

The switching point in the Holt-Laury mechanism allows derivation of the bounds in which a subject's risk parameter must be located. Assume that a subject facing Decision 6 in Table 3.1 chooses Option B for the first time and continues to choose Option B thereafter. This means that for Decision 5 he views:

[5] Holt and Laury used the utility function $U(x)=\frac{x^{1-r}}{1-r}$, their bounds will look different but have similar properties. This was done to allow direct comparison to Cox, Smith and Walker (1988) and Isaac and James (2000).

Table 3.1: Lottery Choices and expected payoff differential of choosing Lottery A over Lottery B

Decision	Option A	Option B	Expected Payoff Differenc
1	1/10 of \$2.00, 9/10 of \$1.60	1/10 of \$3.85, 9/10 of \$0.10	\$1.17
2	2/10 of \$2.00, 8/10 of \$1.60	2/10 of \$3.85, 8/10 of \$0.10	\$0.83
3	3/10 of \$2.00, 7/10 of \$1.60	3/10 of \$3.85, 7/10 of \$0.10	\$0.50
4	4/10 of \$2.00, 6/10 of \$1.60	4/10 of \$3.85, 6/10 of \$0.10	\$0.16
5	5/10 of \$2.00, 5/10 of \$1.60	5/10 of \$3.85, 5/10 of \$0.10	−\$0.18
6	6/10 of \$2.00, 4/10 of \$1.60	6/10 of \$3.85, 4/10 of \$0.10	−\$0.51
7	7/10 of \$2.00, 3/10 of \$1.60	7/10 of \$3.85, 3/10 of \$0.10	−\$0.85
8	8/10 of \$2.00, 2/10 of \$1.60	8/10 of \$3.85, 2/10 of \$0.10	−\$1.18
9	9/10 of \$2.00, 1/10 of \$1.60	9/10 of \$3.85, 1/10 of \$0.10	−\$1.52
10	10/10 of \$2.00, 0/10 of \$1.60	10/10 of \$3.85, 0/10 of \$0.10	−\$1.85

$$EU(A) = 0.5(\$2.00)^r + 0.5\,(\$1.60)^r > 0.5(\$3.85)^r + 0.5\,(\$0.10)^r = EU\,(B) \tag{3.8}$$

and for Decision 6 he views:

$$EU\,(B) = 0.6(\$3.85)^r + 0.4\,(\$0.10)^r > 0.6(\$2.00)^r + 0.4\,(\$1.60)^r = EU(A) \tag{3.9}$$

Solving equation 3.8 and 3.9 shows that a subject who has the utility function in equation 3.7 has a risk parameter that falls in the range $0.59 < r < 0.85$. Table 3.2 lists the risk parameters for possible switches.

Hypotheses

Hypothesis 1: *The aggregated mean CRRAM risk parameter from the first price auction and the mean last switch in the Holt Laury lottery risk parameter are consistent with past*

Table 3.2: Risk parameter derived from last switch

Last Switch	**Range Of Relative Risk Aversion** for $\mathbf{U}(\mathbf{x}) = \mathbf{x}^r$	**Risk Preference Classification**
$1-2$	$r > 1.95$	highly risk loving
3	$1.49 < r < 1.95$	very risk loving
4	$1.15 < r < 1.49$	risk loving
5	$0.85 < r < 1.15$	risk neutral
6	$0.59 < r < 0.85$	slightly risk averse
7	$0.32 < r < 0.59$	risk averse
8	$0.03 < r < 0.32$	very risk averse
9	$-0.37 < r < 0.03$	highly risk averse
10	$r < -0.37$	stay in bed

subject pools, and the mean CRRAM risk parameter in the auction falls within the CRRAM range of the risk parameters derived from the lottery.

Recall that Harrison showed that his subject pool's risk estimate was different than the risk estimate obtained by Cox, Smith and Walker. Isaac and James had the same subject pool participate in two different institutions and found that the risk parameter estimates were not consistent

In my experiment, I have observations for the same subject pool in different institutions. I can also compare my results in the auction to the past results of Cox, Smith and Walker, and Isaac and James, and my results in the lottery to those of Holt and Laury and Prasad and Salmon to see how my subjects' behavior in each institution compares with other subjects in similar institutions. By validating that my subjects behave similarly to past subjects, I can be confident that any further findings in this study are not special to my subject pool. If my subjects as a group behave consistently across institutions, it becomes worthwhile to study them at the individual level. If there was a preference reversal in the mean, the possibility for consistency at the individual level would be eliminated.

Hypothesis 2: *The ranking of subjects by their risk tolerance (both β and r in the* *auction and last switch and r in the lottery) is consistent across the auction and the lottery.*

If subjects are found to have different risk parameter values in different institutions, it could be the case that there is simply a shift and the subjects would maintain their relative ranking i.e. *subject A* would be more risk averse than *subject B* in both institutions. If subjects at least maintain their ranking, we can be confident that the modelling techniques are capturing something salient about the subjects. It could be the case that for some reason context makes the subjects more risk averse in one institution than the other. This would still be troubling when trying to speak generally about behavior, but it would be less troubling than subjects' ranks varying wildly between institutions. If subjects' ranks are not consistent we would have to figure out what kind of semantic game we are playing with the word risk.

Hypothesis 3: *Individual subjects exhibit the same level of risk tolerance in different institutions. Subjects β should be correlated with last switch and under the CRRAM assumptions, a subject's risk parameter in the auction should fall in the corresponding range of risk parameters derived in the lottery.*

The point of developing economic models is to be able to predict how people will behave. If an individual has certain tendencies when making a decision that entails uncertainty, it should be the case that these tendencies are similar regardless of the institution. We can think of it this way, suppose a person does not like peanuts. I can then predict that he will never willingly eat peanuts, and I will feel very confident in my prediction. But, what if this person only dislikes peanuts sometimes? My attempts at predicting his behavior becomes more taxing. I have to start worrying about context, *maybe he likes peanuts when he is at a baseball game, but he does not like peanuts when he is watching baseball at home.* This is a more difficult situation to model. Imagine I had to forecast demand for peanuts. If everyone had consistent peanut preferences, my job is not that hard. But if demand for peanuts changes with context, forecasting is much harder. Now I have to approximate baseball attendance. I can still do a forecast, but I will be much less confident in my prediction. If we think about risk aversion, we can view it as someone not liking risk. As an economist, I will feel much more comfortable about my predictions if people do not like risk by the same amount whether or not they are at a baseball game.

3.4 Results

Result 1 : *Subjects (as a group) in the first price auction behave consistently with the subjects in Cox, Smith and Walker. Subjects in the Holt-Laury lottery behave consistently with the subjects in Holt and Laury and Prasad and Salmon. In the current study, the mean risk parameter in the auction, calculated with the CRRA model, falls within the range that these same subjects exhibited in the lottery.*

In this section I will show that my subjects behaved as other subjects did in the same institution. By replicating the past experiments, we can diminish any critique that further results are determined by a subject pool effect. Table 3.3 shows the last switch observed for subjects using the H-L mechanism in my experiment and two past experiments. Holt and Laury report the mean last switch in their experiment is 6.2 which implies risk aversion, and a CRRAM risk parameter in the range $0.59 < r < 0.85$. Prasad and Salmon report that the mean last switch in their experiment is 7.06 which implies a risk parameter in the range $0.32 < r < 0.59$. I find that the mean last switch in my experiment is 6.86,[6] which implies a risk parameter in the range $0.32 < r < 0.59$. The reported last switches for all three experiments are not statistically different ($p - value = 0.63$ for comparison with P-S, $p - value = 0.11$ for comparison with H-L). Subjects in my experiment are behaving consistently with subjects in prior experiments.

Table 3.4 shows the mean implied risk parameters of subjects participating in a first-price, sealed bid auction. I find a mean estimate of $r = 0.37$ which is between the parameters found by Cox, Smith and Walker and Isaac and James. CSW find a mean estimate of $r = 0.35$, I-J find a mean estimate of $r = 0.50$. All of these estimates imply risk aversion. My results are not statistically different than CSW ($p - value = 0.6624$). However, my results are statistically different than I-J ($p - value = 0.0162$).[7] The key aspect of risk preference is not changed, just the level. This could easily be due to my subject pool being slightly different than I-J. It should be pointed out that my experiment and CSW had four human

[6]Only subjects that conform to the Nash equilibrium linear bid function, ($\alpha = 0$) in the first price auction, are used for comparison. The mean last switch point for all subjects that is 7.03. This result is not statistically different than 6.86 ($p - value = 0.683$).

[7]I-J state that their results are not statistically different from CSW. CSW have a smaller subject pool than I do, and therefore a larger confidence interval. I-J's estimate falls within CSW's interval, but not mine.

Table 3.3: Mean last switch in lottery with different subject pools

Lottery	**H-L**	**P-S**	**E**
Last Switch	6.2	7.06	6.86
Derived risk parameter	$0.59 < r < 0.85$	$0.32 < r < 0.59$	$0.32 < r < 0.59$

Table 3.4: Mean risk parameter in auction with different subject pools

Auction	**CSW**	**I-J**	**E**
Risk Parameter (r)	0.35	0.50	0.37

Table 3.5: Mean risk parameter in auction and lottery with the same subject pool

	Auction	**Lottery**
Risk Parameter (r)	0.37	0.32-0.59

subjects bidding against each other, where I-J had one human bidding against four risk neutral robots, though Walker, Smith and Cox (1987) found no significant difference in how subjects behave when bidding against robots instead of humans.

Table 3.5 shows that the implied risk parameter for the group in the auction falls within the range of the implied risk parameter for the group in the lottery. This does not yet prove that subjects at the individual level behave the same way in each institution, but allows for the possibility. There would be no need to run individual subject level tests if the result had been the opposite, and we could be confident that subjects do not behave the same way in different institutions.

Result 2: *The ranking of subjects from most to least risk averse is consistent across* *institutions.*

The first way I have chosen to test cross task consistency is to rank subjects in each institution and examine if these rankings are consistent across institutions. I do not need to assume an underlying model to do this because subjects can be ranked based on their β and their *last switch.* After ranking the subjects from most risk averse (high values of β) to least risk averse (low values of β) a Spearman test was conducted to test rank preservation. I find that there is a positive correlation, ($\rho = 0.35$) but it is not significant at the 10% level (it is significant at the 11% level). So we do see a positive correlation in ranks, but the significance is tenuous. This lack of significance could be due to the way subjects' ranks are determined in the auction versus how they are ranked in the lottery. Recall that in the auction, there is an explicit estimate of β. In the lottery, subjects' ranks are determined by where they switch to *Option B.* The lottery does not give a point estimate for risk aversion. This could possibly create a problem because tiny differences in the auction could lead to a large difference of where a subject ranks in the auction, but subjects who last switch at 7 cannot be ranked within that bin. It is easier to think about this in terms of the CRRAM. With the CRRAM we can derive point estimates in the auction, but only bins over a range of risk parameters in the lottery. For example, subjects 1,2, and 19 in the lottery switch to Option B at decision 7 which places their risk parameter in the range $0.32 < r < 0.59$. In the auction, the subjects have $r = 0.42$, 0.58, and 0.36 (respectively). This means that subjects 1,2, and 9's estimates from the auction fall within the range found in the lottery. But, when these subjects are ranked, the lottery ranks them all at 11 (since they are tied), while the auction ranks them at 15, 19, and 12. I have taken subjects' risk parameter estimates from the auction and put them into the same range that the estimates from the lottery fall (Table 3.6). *What I am forcing here is not that ranks across institutions be similar, but that ties in one institution are considered ties in the other institution.* Figure 3.1 shows a scatterplot of both the actual rankings and the hypothetical rankings from the auction on the rankings from the lottery. Using these created last switch estimates, I have re-estimated Spearman's ρ and find a much stronger correlation ($\rho = 0.59$) with significance at the 5% level.

Result 3: *Individual subjects behave consistently across different institutions. Subjects' risk aversion measured by both their* β, and their *CRRAM parameter* (r) *in the first price auction is consistent with subjects' risk aversion measured by their last switch in the lottery.*

Table 3.6: Subjects revised auction rankings. This places subjects' auction risk parameters in bins consistent with those derived in the lottery.

Subject	r (auction)	r (lottery)	Rank (auction)	Rank (lottery)	Rank (new)
1	0.42	$0.32 < r < 0.59$	15	11	16
2	0.58	$0.32 < r < 0.59$	19	11	16
19	0.36	$0.32 < r < 0.59$	12	11	16

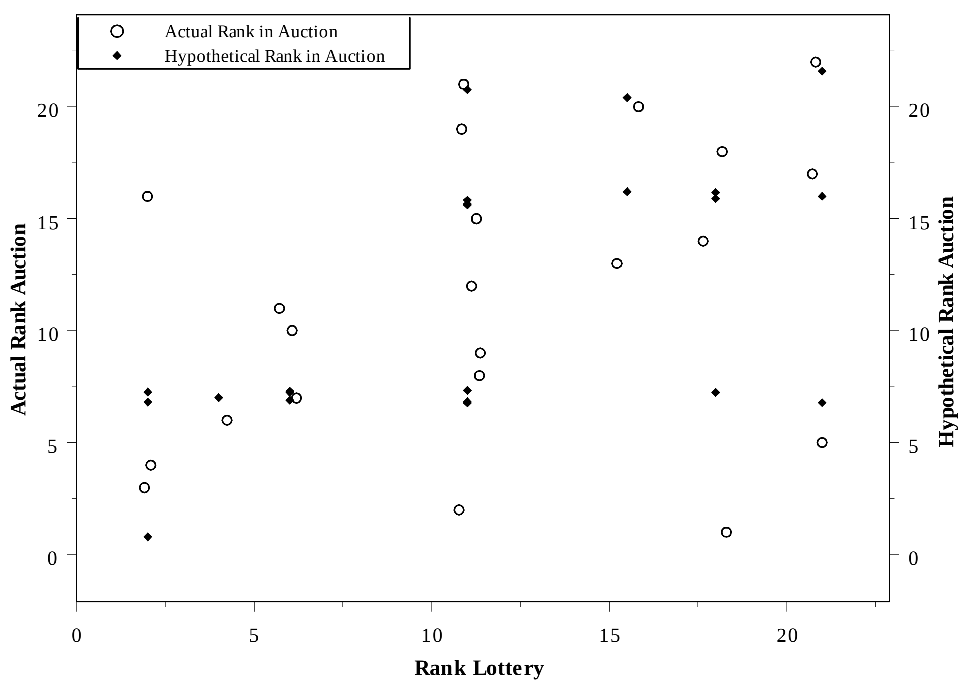

Figure 3.1: Within subject comparison between institutions (ranks): Actual rank is represented by circles, Hypothetical by diamonds

Further, the CRRAM parameter from the first price auction is consistent with the CRRAM parameters derived from the lottery.

We saw in Result 1 that my subjects behave as subjects in past experiments behaved. I can now address the more important question of whether individuals behave similarly across institutions. The first test I will use is an Ordered Logit regression. This test is used because the H-L Lottery does not allow for derivation of point estimates for a subject's risk preference. The Ordered Logit allows for using the *last switch* levels directly as the dependent variable. Table 3.7 reports the results of an Ordered Logit Regression where I have regressed the *last switch* subjects made in the H-L Lottery on the coefficient of a subjects value estimated by their linear bid function in the first price auction (β). Table3.7 also reports the results of the same test using the CRRAM derived risk parameter (r) instead of β. I find that there is a significant relationship between the two variables for both β and *last switch*, and r and *last switch*. In other words, the more risk averse a subject is in the first price auction, the more likely it is that he would be more risk averse (have a high switching point) in the H-L Lottery. Given the intercept (α_1 through α_6)[8] and the coefficient, it is possible to construct probabilities of making a specific last switch given some β. For example, to calculate the probability that an individual with $\beta = 0.8$ makes his last switch at decision 7:

$$\frac{e^{\alpha_4 - \beta(0.8)}}{1 + e^{\alpha_4 - \beta(0.8)}} = \frac{e^{15.482 - 16.433(0.8)}}{1 + e^{15.482 - 16.433(0.8)}} = 0.912 \tag{3.10}$$

Figure 3.2 shows the estimated probability of a subject having a particular Last Switch given their estimated β. For example, if a subject on average bids 80% of his value ($\beta = 0.8$) there is a 91.2% chance that he made his last switch at decision 7 or lower, and more specifically, there is a 22% chance that he made his Last Switch exactly decision 7. In the first panel when $\beta = 0.6$ we can see that the probability is heavily skewed towards a subject making their last switch early, while when $\beta = 0.99$ the probability is heavily skewed towards of a subject making their last switch late. We can see similar results using the risk parameter r in Figure 3.3. The data is limited to having a Last Switch lower bound of 4 because no

[8] In an orderd logit, there is an intercept for each section of the dependent variable. In this case, subjects were observed to make a last switch as early as decision 4 or as late as decision 10. There are 6 intercepts. There need not be an intercept for last switch 10, as the probability for any risk parameter being last switch 10 or lower is 1.

Table 3.7: Ordered Logit Regression of the subjects' last switch in the lottery on the beta coefficient in the auction and the CRRAM risk parameter

	Coef	**Std Err**	**P>|z|**		**Coef**	**Std Err**	**P>|z|**
Beta Auction	16.433	8.224	0.046	**Risk Auction**	-4.199	2.06	0.042
α_1	12.523	7.158	0.080	α_1	-3.7285	1.18	0.001
α_2	13.493	7.220	0.062	α_2	-2.738	1.028	0.008
α_3	13.956	7.248	0.054	α_3	-2.27	0.983	0.021
α_4	15.482	7.393	0.037	α_4	-0.742	0.866	0.391
α_5	16.322	7.460	0.029	α_5	0.0948	0.875	0.914
α_6	16.700	7.489	0.026	α_6	0.4692	0.905	0.605

subjects were observed to switch earlier than that.

The results of the Ordered Logit are consistent with an individual behaving similarly in both institution. However, when we are examining the CRRAM model, the Ordered Logit is only testing that the auction derived parameter is consistent with the dependent variable *last switch*. So, a higher risk parameter was consistent with a lower *last switch* but it did not test if the numerical value of the risk parameter fell within the range of numerical values that can be backed out of the lottery. We can better evaluate the CRRAM results by looking explicitly at each subjects derived parameter in both institutions. Table 3.4 shows subjects derived risk parameters in both the lottery and the auction. There are 7 out of 22 subjects (32%) whose risk parameter, derived from the auction, falls within the range of the risk parameter derived in the lottery. Of the remaining 15 subjects, 10 are more risk averse in the first price auction than in the lottery institution. This means that 17 out of the 22 (77%) are at least as risk averse in the lottery institution as they are in the first price auction. It is noteworthy that three of the subjects who were more risk averse in the lottery than the auction did not switch to *Option B* in the lottery until the high payoff was assured. These three subjects were quite risk averse in the auction ($r = 0.11, 0.18, 0.44$). This is of interest because if a subject bids very close to their value (extremely risk averse) the lowest r that can be estimated for them is 0. Whereas subjects who do not switch to *Option B* in the lottery until it a sure thing will have an estimated risk parameter of $r < -0.37$. So it would

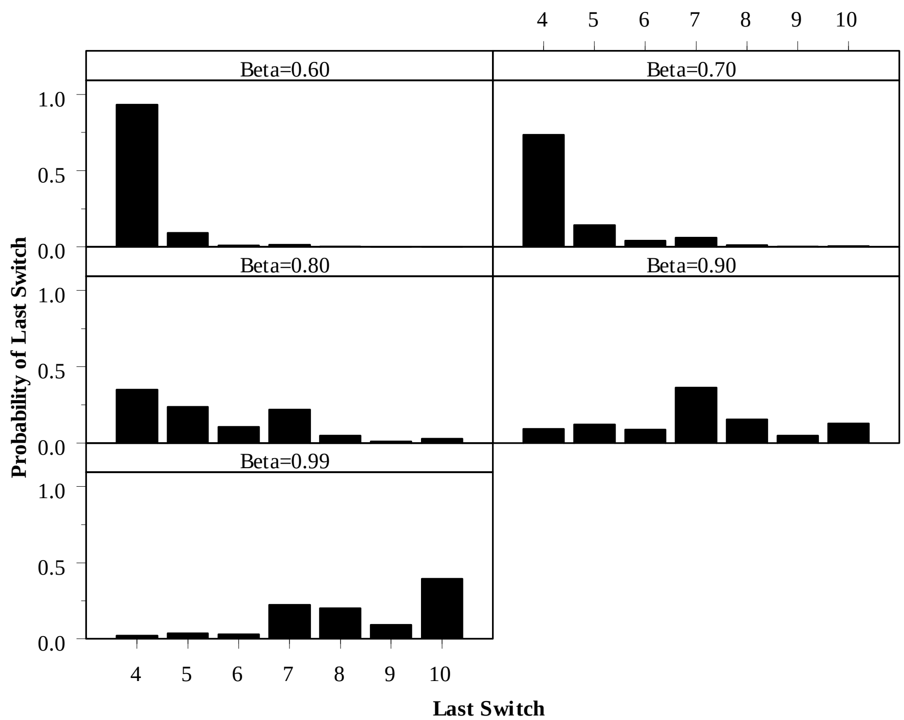

Figure 3.2: Probablility of specific last switch for a given Beta

be impossible for a subject to be as risk averse in the auction as someone who made their *last switch* at decision 10. If we look at Figure 3.4 (a plot of risk parameter ranges from the lottery plotted on the risk parameter values derived from the auction) we can see that there appears to be a strong positive correlation between the range from the lottery and the point estimates from the auction.

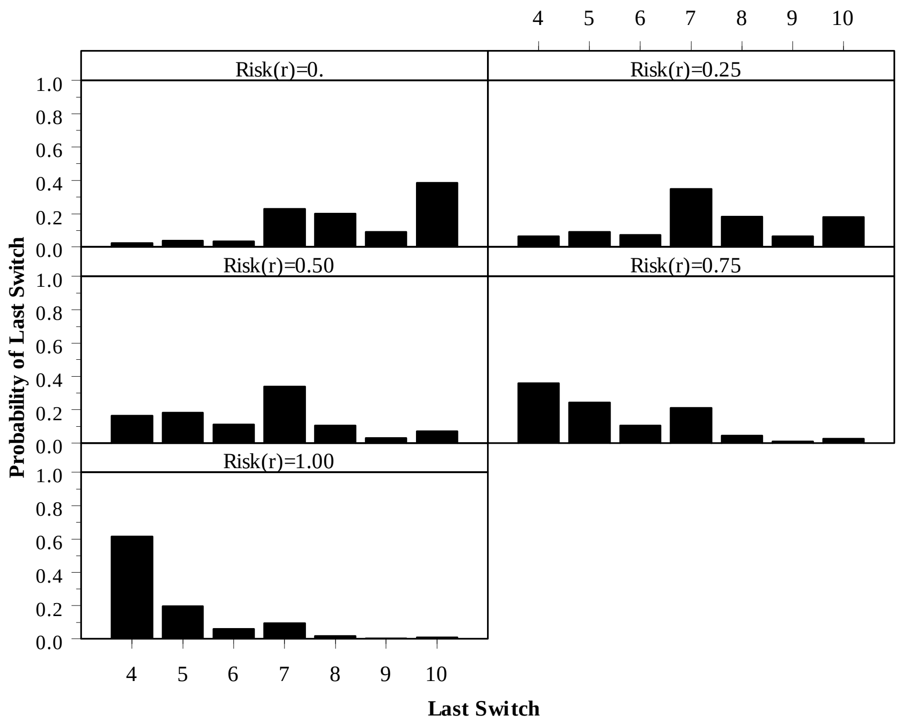

Figure 3.3: Probablility of specific last switch for a given risk parameter

Subject	RP (HL)	RP (FPA)	HL-FPA		Subject	RP (HL)	RP (FPA)	HL-FP
1	$0.32 < r < 0.59$	0.42	0		16	$0.32 < r < 0.59$	0.29	0.03
2	$0.32 < r < 0.59$	0.58	0		19	$0.32 < r < 0.59$	0.36	0
3	$0.32 < r < 0.59$	0.09	0.23		20	$r < -0.37$	0.18	-0.55
4	$1.15 < r < 1.49$	0.45	0.70		21	$0.85 < r < 1.15$	0.06	0.79
5	$0.32 < r < 0.59$	0.30	.02		22	$0.85 < r < 1.15$	0.55	0.30
6	$0.85 < r < 1.15$	0.40	.45		24	$-0.37 < r < 0.03$	0.22	-0.19
9	$0.32 < r < 0.59$	0.67	-0.08		26	$0.03 < r < 0.33$	0.26	0
10	$0.32 < r < 0.59$	0.19	0.13		28	$0.03 < r < 0.33$	0.30	0
12	$0.59 < r < 0.85$	0.39	0.20		29	$0.03 < r < 0.33$	0.30	0
13	$r < -0.37$	0.11	-0.48		30	$1.15 < r < 1.49$	1.09	0.06
15	$r < -0.37$	0.44	-0.81		31	$0.59 < r < 0.85$	0.59	0

Table 3.8: Subject CRRAM risk parameter estimates in each institution

Correlation, however, is not enough to say that subjects behave the same way in each institution. It could be the case that subject are always more risk averse in one institution than the other. We have seen in Result 2 that subjects are ranked similarly in both institutions, but that does not guarantee that the risk parameter for the subject is the same in both institutions. To examine this, I first have to see if there is, in fact, significant correlation. To do this, I have used subjects' r estimates from the auction to see where they would have last switched in the lottery from *Option A* to *Option B* (i.e. assume the risk parameter from auction is true and then derive how a subject with that risk parameter would act in the lottery). I call this the *implicit last switch.* We can see the result of this in Table 3.9, and Figure 3.5 shows a scatterplot of the *actual last switch* on the *implicit last switch.* With this sorting, I can now check for correlation between the *actual last Switch* and the *implicit last switch.* I find that these are positively correlated at the 5% level (Pearson $r = 0.56$). With the assurance of this correlation, I can now examine the relationship between the two institutions by regressing the *actual last switch* on the *implicit last switch.* Table 3.10 reports the results of this test. The coefficient of the *implicit last switch* is positive and significant at the 1% level. The result of the coefficient being equal to 1.16 is somewhat counter-intuitive given that the constant term is not significantly different than zero, since this implies that subjects are slightly more risk averse in the lottery than the auction. This is counter-intuitive because we have seen that 10 of the 15 (67%) subjects whose derived risk parameter from the auction was outside of the bounds of that derived by the lottery were more risk averse in the auction than in the lottery. The constant term is not reported as being significantly different than zero, but it is reported as negative. I have suppressed the constant term and re-run the regression and find that forcing the constant term to zero returns the more intuitive result of the coefficient on the *implicit last switch* being less than 1 (the coefficient is equal to 0.93). The most important result is that the coefficient is close to one. This means that not only is the behavior correlated, but subjects are behaving the same way in each institution.

Table 3.9: Subjects last switch in lottery and implicit last switch derived from the auction

Subject	LS (actual)	LS (implicit)	Subject	LS (actual)	LS (implicit)
1	7	7	16	7	8
2	7	7	19	7	7
3	7	8	20	10	8
4	4	7	21	5	8
5	7	8	22	5	7
6	5	7	24	9	8
9	7	6	26	8	8
10	4	8	28	8	8
12	6	7	29	8	8
13	10	8	30	4	5
15	10	9	31	6	6

3.5 Conclusion

When developing economic models, certain assumptions are made. One assumption commonly made is that individuals will respond to uncertainty in a consistent way. Individuals may be different, but there is internal consistency to their behavior (Joe is not the same as John, but Joe behaves the same way in different institutions). Recent laboratory experiments have questioned the validity of this assumption. I add to this literature by having the same subjects participate in two different institutions (the Holt and Laury lottery and a first price auction) that allow for inferences to be made about their risk tolerance. A nice feature of the two institutions I have used is that they do not rely on specific utility functions for analysis of risk preferences. Though it is possible to back out specific parameters if they are desired.

Subjects in my experiment behave consistently with subjects in past experiments given the same institution. So, there is nothing special about my subject pool in this regard. Further the behavior of the group across institutions is similar. The most important result is that individuals behave consistently across these two institutions. Using the non-parametric measures, I find that subjects who are more risk averse in one institution are likely to be more risk averse in the other institution. Beyond that, I find that the parametric estimates

Table 3.10: Results of regressing the actual last switch on the implicit last switch, and results of same regression with a suppressed constant

	Coef.	**Std. Err.**	**t**	**P>|t|**
Last Switch (implicit)	1.16	0.38	3.04	0.006
Constant	-1.79	2.86	-0.63	0.539

	Coef.	**Std. Err.**	**t**	**P>|t|**
Last Switch (implicit)	0.93	0.045	20.69	0.000
Constant	-	-	-	-

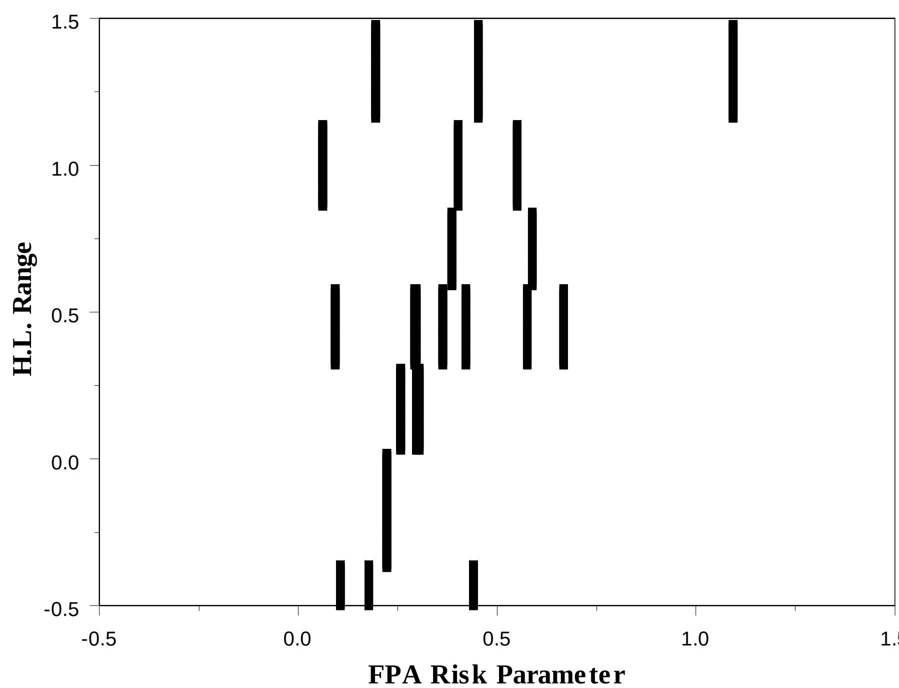

Figure 3.4: Scatterplot of derived CRRAM risk parameter from the lottery and the auction

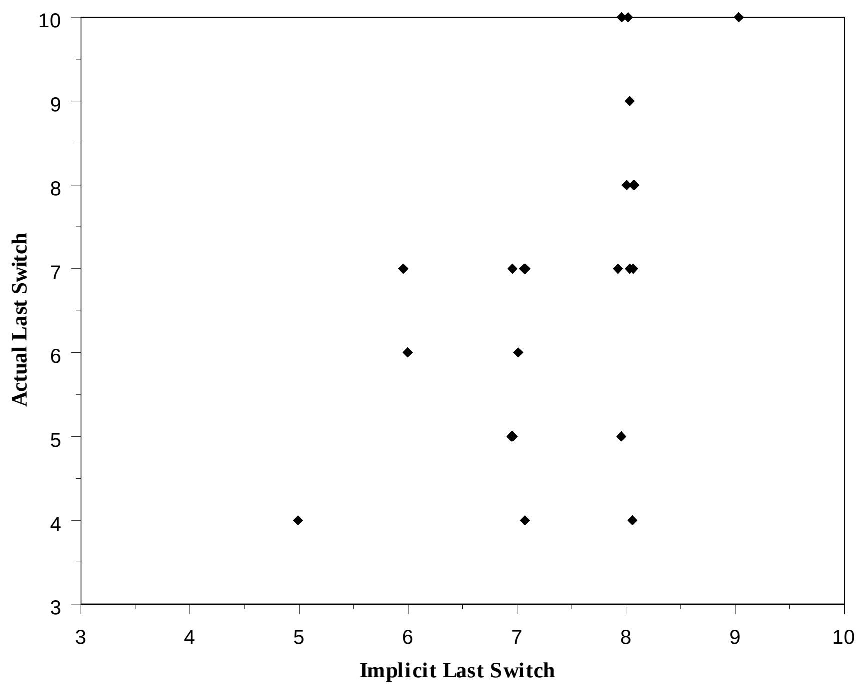

Figure 3.5: Scatterplot of the actual last switch made by subjects on implicit last switch derived by subject behavior in the auction

derived by using the CRRA model hold up quite well across institutions.

These results do not allow me to generalize that individuals will always behave with consistent risk preferences in all institutions. There is still the question of what was happening in the BDM mechanism in past research. Isaac and James (2000) is titled "Just Who Are You Calling Risk Averse?," and there is a question in the literature of whether subjects in first price auctions merely bid 'as if' risk averse. My results support the argument that subjects bid 'as if' risk averse because they are, in fact, risk averse.

3.6 Appendix

3.6.1 Experimental Instructions

Experiment Instructions I

In each round of this series you will be asked to make a choice between two lotteries that will be labeled A and B. There will be a total of 10 rounds and after you have made your choice for all 10 rounds, one of those rounds will be randomly chosen to be played. Lottery A will always give you the chance of winning a prize of \$2.00 or \$1.60, while lottery B will give you the chance of winning \$3.85 or \$0.10. Each decision round will involve changing the probabilities of your winning the prizes. For example in round 1, your decision will be represented on the screen in front of you (activate program now):

Your decision is between these two lotteries:

Lottery A: A random number will be drawn between 1 and 100. You will win

\$1.60 if the number is between 1-90 (90 % chance)

\$2.00 if the number is between 91 and 100 (10 % chance)

Lottery B: A random number will be drawn between 1 and 100. You will win

\$0.10 if the number is between 1 and 90 (90% chance)

\$3.85 if the number is between 91 and 100 (10% chance)

If you were to choose lottery B and this turns out to be the round actually played, then the computer will generate a random integer between 1 and 100 with all numbers being equally likely. If the number drawn is between 1 and 90, then you would win \$0.10 while if the number is between 91 and 100, then you would win \$3.85. Had you chosen lottery A then if the number drawn were between 1 and 90 you would win \$1.60 while a number between 91 and 100 would earn you \$2.00.

All of the other 9 choices will be represented in a similar manner. Each will give you the probability of winning each prize as well as translate that probability into the numerical range the random number has to be in for you to win that prize.

At the end of the 10 choice rounds, you will be asked to press a button that will allow the computer to determine your payment. When you do so, the computer will randomly pick one of the 10 rounds to base your payment on, remind you of the choice you made in that round and draw the random number between 1 and 100 to determine your earnings.

Are there any questions before you begin making your decisions?

We ask that you follow the rules of the experiment and in particular we again ask that you do not talk or look at the screens of other participants during the experiment. Anyone who violates the rules may be asked to leave the experiment with only the $10.00 show-up fee.

You will now start the sequence of 10 choices. You will be able to go through the choices at your own pace, but we will not be able to continue the experiment until everyone has completed this series.

So now please look at your computer screen, you can determine your choice by clicking on the circle beside it. Now you may begin making your choices. Please do not talk to anyone while you are doing this; raise your hand if you have a question.

Experiment Instructions II

In this part of the experiment, you will be a bidder in an auction. You are one member in a group of four. On your handout you will see something called "your value" this is how much the item that you are bidding on is worth to you, the other three players have their own value. You have to decide how much you are willing to bid for this object, given your value. You will be bidding against 3 other people in this room. The person who submits the highest bid (on the handout you will see a box marked "your bid") will be the winner of the auction. The winner of the auction will earn their value minus their bid. For example, assume your value is 65, and you bid 55. If 55 is the highest bid, your profit would be 65-55=10.

You will see, on the second page of your handout, what the winning bidder's screen will look like at the end of a round. Notice that "winning bid" is equal to "your bid." If more than one person submits the same bid, and this is the winning bid, the computer will randomly select a winner from those that were tied. On this handout, “your value” is 53, and “your bid” is 45, therefore “your earnings” for the round are 53 – 45 = 8 ECUs (100 ECUs = $2.00).

The computer procedure used to get “your value,” and the values of the other members of your group, mimics drawing numbered balls from a container. For each round, a container would contain 100 balls uniquely numbered with the numbers 1,2, and 3, and so on through 100. This means the highest value anyone in your group could have is 100 and the lowest value is 1. Your value gives you no information about the other group members’ values.

There will be 20 rounds for this portion of the experiment

Experiment Instructions II

In this part of the experiment, you will be a bidder in an auction. You are one member in a group of four. On your handout you will see something called “your value” this is how much the item that you are bidding on is worth to you, the other three players have their own value. Your value will be a random integer between 0 and 100, with all numbers being equally likely. “Your value” is randomly selected for each member of your group. It is highly likely that all the members in your group all have different values.

You have to decide how much you are willing to bid for the object, given your value. You will be bidding against 3 other people in this room. The person who submits the highest bid (on the handout you will see a box marked “your bid”) will be the winner of the auction. The winner of the auction will earn their value minus their bid. For example, assume your value is 65, and you bid 55. If 55 is the highest bid, your profit would be 65-55=10. If you do not win, you earn zero.

You will start this portion of the experiment with 25 ECUs. If you bid above your value (i.e. “your bid” is greater than “your value”), and you win the item, you will make losses, and they will be subtracted form your initial 25 ECUs. You are not prohibited from bidding above your value, but you can ensure that you never make losses as long as you never bid above your value.

You will see, on the second page of your handout, what the winning bidder's screen will look like at the end of a round. Notice that "winning bid" is equal to "your bid." If more than one person submits the same bid, and this is the winning bid, the computer will randomly select a winner from those that were tied. On this handout, "your value" is 53, and "your bid" is 45, therefore "your earnings" for the round are 53 – 45 = 8 ECUs (100 ECUs = $4.00).

There will be 30 rounds for this portion of the experiment

CHAPTER 4

Why Work When You Can Shirk? Worker Productivity in an Experimental Setting

4.1 Introduction

There is a long standing and mature theoretical literature dealing with the principal-agent problem (Mirrlees (1975), Grossman and Hart (1983)). A key aspect of the principal-agent problem is the motivation of workers who are rational cheaters. Nagin (2002) defines a rational cheater as someone who will shirk when the marginal benefit of doing so exceeds the marginal cost. A rational cheater will exert low effort on the job if he thinks he can get away with it. Recent reports estimate that shirking workers cost employers billions of dollars in productivity losses yearly[1]. Employers (principals) who are aware of the financial incentives they are giving employees (agents) may introduce a monitoring system with performance goals to alleviate the perceived problem. These issues with worker motivation are difficult and complex. Consider a recent New York City court case:

> On March 9, 2006, John B. Sooner, a New York City administrative law judge, recommended that Toquir Choudhri, a 14-year veteran of the city Department of Education, receive only a reprimand for disobedience, even though supervisors wanted him fired for using the Internet for personal matters[2]. Spooner wrote that Choudhri credibly stated that he completed all assignments given to him

[1] Frauenheim, Ed. "Stop Reading This Headline and Get Back to Work." *CNET,* Monday, July 11, 2005, http://news.com.com/Stop+reading+this+headline+and+get+back+to+work/2100-1022_3-5783552.html

[2] Klopott, Freeman. "Should You Be Fired for Using the Internet While at Work?" *PC World,* Tuesday, May 02, 2006, http://pcworld.com/article/id,125597-page,1/article.html?RSS=RSS

> by his boss and used the internet while he awaited further assignments. These statements were corroborated by the absence of proof that Choudhri was ever criticized for poor productivity or for not completing specific assignments.[2] The New York City Chancellor of Education, Joel Klein, decided to fire Choudhri anyway. Klein stated that 'the penalty of termination is appropriate and not shocking to one's sense of fairness, ... Choudhri's abuse of the Internet at the time he is supposed to be performing his job demonstrates his disinterest in the job.'[3]

The worker in the above case was fired for shirking on the job when his employer found him surfing the internet. The worker did not think he deserved to be fired because he had completed all of his assignments. The worker thought he was being monitored in regard to fulfillment of some quota, and he had fulfilled his quota, but the employer disagreed. This case demonstrates the problems caused when the monitoring system is not well delineated, but it also shows how concerned some employers are about any behavior consistent with shirking. In this paper, I will examine how well these types of schemes work when they are clearly stated and consistently enforced. I will contrast my results with previous research which showed that workers may work harder than required by explicit financial incentives.

Past experimental research (Cadsby et al. (2004), Dickinson and Villeval (2005)) has shown that some workers in a laboratory setting work without incentives. It is important to investigate if these laboratory results indicate behaviors we would see in a real work setting, or if the possibility exists that these observations are an artifact of the experimental design. This paper attempts to place subjects in a more refined laboratory setting to get a cleaner look at subject behavior towards work effort with low financial incentives. Cadsby et al. (2004), and Dickinson and Villeval (2005) were not specifically looking at effort with low incentives, so there is much remaining value to their research even if the observations of subjects working without incentives are an artifact of their design.

[2] Department of Education v. Choudhri, New York City Office of Administrative Trials and Hearings, No. 722/06 (3/9/06)

[3] Associated Press. "NYC Fires Man For Web Surfing At Work." *CBS News,* http://www.cbsnews.com/stories/2006/05/06/tech/main1596034.shtml

A common approach used in the laboratory to investigate the principal-agent problem is to give subjects a cost function and have them choose some effort level. Nalbantian and Schotter (1997), for example presents an experiment in which a subject is monitored with probability. If the subject is not expending a certain level of 'effort', he will be terminated. The 'effort' in this case is not physical exertion but rather a figurative effort. This number they pick is indeed interpreted as effort and therefore has the property that effort is now explicit. While this matches clearly with their models, it is not clear that subjects perceive this choice as analogous to physical or mental exertion. The (e) chosen by the subject is costly to the subject but it is possible that this is too abstract to model real work. Putting the workers through a real effort experiment will allow an answer as to whether simply choosing effort garners the same behavior as exerting effort, and if it does not, one can be confident that the real effort experiment is a better proxy for the workplace.

While an agent's outside option can be represented rather easily in a theoretical framework, it is not so trivial to do in a laboratory. To think of this one must place himself in the position of the subject in the experiment. The subject arrives at the laboratory and is assigned a computer terminal. He is given the option of engaging some task and earning X or he can take the outside option (sit still) and earn Y. If the disparity between the outside option and the participation option is not large, there exists the possibility that the subject will engage the game to avoid boredom.

The aforementioned research by Cadsby et al. (2004), and Dickinson and Villeval (2005) found that subjects contribute effort even when they have no financial reason to do so. The idea of a moral imperative not to shirk is given as a reason for this behavior by Dickinson. Cadsby's experiment gives subjects seven scrambled letters and the subjects are instructed to make as many words as possible in a given time period. The subjects are allowed to choose a piece rate or a flat rate scheme, the piece rate will pay them per word, and the flat rate will give them some stated amount with no requirement on word creation. These incentives should lead students who think they are endowed with word creation ability and low effort cost to choose the piece rate scheme and the subjects who have high effort cost or are not good at word creation to choose the flat rate scheme. The interesting observation is that there are significantly more than zero words created by the subjects who choose the flat rate scheme. This is contrary to the incentive structure, and one would not believe the students feel a moral imperative to unscramble letters. Dickinson and Villeval (2005) use a real effort task and they observe that some of the subjects (25%) contribute at or above the desired output level even when monitoring is set to zero. Dickinson, as noted earlier, suggests this as either intrinsic motivation or integrity and commitment to moral principles. It is possible that subjects feel a moral imperative because they are interacting with a human principal, but that is not the case in Cadsby (2004). The intrinsic motivation argument is plausible in both experiments. The subjects might enjoy unscrambling letters, and they might enjoy moving along a curve to get a high value, but it is not clear that observations made under these conditions should be interpreted as analogous to intrinsic motivation one might experience in the real world. In both experiments, the subjects could engage in effort, or do nothing. It is possible that they were engaging the tasks because they were bored. This is similar to Choudri's claim that he browsed the internet only because he had no other work to do. If the subjects in the experiment do not engage the task, they have nothing else to do.

Other experiments have shown that outside options have important effects. Lei, Plott, and Noussair (2001) show that excess trading in an asset market can be reduced by giving subjects something to do beyond trading in the asset market. Pevnitskaya and Palfrey (2004) show that over entry into an auction can be reduced by allowing subjects an outside option of a computerized version of rock-paper-scissors. Van Dijk et al. (2001) conduct an experiment where subjects are enabled to work on two tasks in the same period. The earnings from one task go in to a group account while the earnings from the other task go in to a private account, similar to a public goods game. This alleviates the effort only due to boredom problem that is plausible in the Cadsby(2004) and Dickinson (2004) papers. The specific task the paper uses has the subjects search a grid looking for the highest payoff. The idea of having two of these for the subjects to play cures the boredom critique.

The experimental design for this paper allows for subjects to play a valuable outside option. This should eliminate play in the primary task when there is no financial incentive to play the primary task. Then the question of how much monitoring is necessary to get the desired amount of effort can be addressed. The desired amount of effort will be explicit in terms of a quota.

Table 4.1: Task A

Probability of being monitored:	25%
Quota:	2
Number Completed:	2
Type this:	7072
Enter here:	
You entered:	

The paper proceeds as follows: Section 4.2 contains the experimental design, section 4.3 contains a simple model of predicted behavior in the experiment, and section 4.4 reports analysis of the results of the experiment. Finally, there will be some concluding remarks.

4.2 Experimental Design

The experiment conducted for this paper was designed to give subjects incentives similar to those faced by many workers. There are 30 periods in the experiment, and each period lasts 45 seconds with 30 seconds between each period.. The experiment consists of a primary task (Task A) and a secondary task (Task B). The subjects can choose to split effort, at their discretion, between the two tasks. Task A is designed to mimic work for an employer. The subject will view a randomly generated four-digit number and his task is to type the same number in the space provided. Every time a subject has completed typing the number, he can hit a button and start on a new randomly generated number. An example of this can be seen in Table 4.1. The payoff to the subjects depends on the quota they face and a monitoring level. The subjects earn an effective wage of 300 ECUs per period. If they are monitored and the amount of four-digit numbers typed is less than the quota they are considered fired without pay, and their earnings are zero. If a subject is not monitored, he earns 300 ECUs from Task A regardless of whether or not his quota is met.

The other task subjects can engage in (Task B) is a matching pennies game, shown in Table 4.2, that they will play against a computer opponent. The computer is playing a Mixed-Strategy Nash Equilibrium of choosing heads 50% of the time and tails 50% of the time. The matching pennies game was chosen because it is a cognitively easy task, and the chosen payoffs have the property that the expected utility is increasing in the amount of

Table 4.2: Matching Pennies

		Computer Heads	Computer Tails
Subject	Heads	2,0	0,2
	Tails	0,2	2,0

time spent on the task. This is intended to mimic the utility a subject would receive from shirking, whether it be reading the newspaper, talking to friends, or browsing the internet. Without Task B, it is possible that the subject would play Task A simply because they are bored and find Task A more stimulating than sitting quietly. With Task B, the subjects could still choose to sit still, but they now have the alternative action involving activity of an outside option that they can earn money by playing.

Any earnings from Task B are in addition to the 300 ECUs earned in Task A, but they are not guaranteed. If the subject is monitored in task 1, and their number completed is less than their quota, any earnings they made in Task B are wiped out. This is a punishment akin to getting fired. One could argue that the subject should keep their outside earnings since an employer can fire an employee without pay, but cannot take away the utility they received from shirking. There will be some fine associated with getting fired and for simplicity the amount of that fine will be equal to the earnings form Task B. This is done to exact some punishment beyond lost wages. In the laboratory, the worker will be rehired the next period, so just taking away the wages for one period is not punitive enough. Regardless of how much is potentially earned on Task B, a subject who is monitored and has not met his quota always earns 0 ECUs for the period.

In order to observe the subjects responses to a variety of situations, there were 6 different monitoring levels (0%, 15%, 25%, 50%, 75%, and 100%) and 5 different quotas (8, 12, 15, 18[3], and 25). This led to 30 combinations of monitoring levels and quotas and each subject faced each combination once. This resulted in the 30 periods mentioned before. The order that the subjects saw the various quotas and monitoring levels was determined randomly

[3]The quota level of 18 only applies to subjects 9-32. The first 8 subjects had a quota of three instead of 18.

prior to the experiment. At the end of every period the subjects would see a screen that would inform them whether or not they were monitored, and their payoff at the end of each period.

The experiment was conducted over four sessions and 32 subjects participated. The subjects for the experiment were undergraduate students at the Florida State University. The experiment was computer based and conducted with z-Tree software (Fischbacher 1999). When the subjects arrived, they were assigned to computer terminals. The instructions for the experiment (see Appendix) were read aloud and subjects had a chance to ask questions. The subjects were paid a \$10.00 show up fee, and were able to earn money based upon their performance in the experiment. The average earnings per subject for the experiment were \$24.96.

4.3 Theory

Given the incentive structure in the experiment, we can now develop a model to predict the behavior of the subjects. Let t be the amount of time a subject spends on Task A, let $m \in (0, 1)$ be the monitoring level the subject faces at the beginning of each period, let $Q \in (3, 25)$ be the quota the subject faces at the beginning of each period. Let $\phi_i(t)$ be the payoff a subject receives for working on Task B ($\frac{d\phi_i}{dt} < 0$ because the more time a subject spends on Task A, the less time they have for Task B). The probability that a subject meets the quota (Q) in time (t) is represented by the term $P_i(t, Q)$. This probability is increasing in t $\left(\frac{dP_i}{dt} > 0\right)$, because the more time spent on Task A, the more likely they are to meet the quota; $P_i(t, Q)$ is decreasing in Q $\left(\frac{dP_i}{dQ} < 0\right)$ because the higher the quota, the more difficult it is to meet it given the 45 second time constraint. Both $\phi_i(t)$, and $P_i(t, Q)$ are individual specific due to the heterogeneity in subject ability (subjects who are good at Task A could also be good at Task B, but that is irrelevant to the analysis of this problem). Recall that subjects can earn 300 ECUs each period if they meet the quota, or are not monitored. The payoff function subjects face can be represented as:

$$\Pi_i(t, Q, m) = mP_i(t, Q)(300 + \phi_i(t)) + (1 - m)(300 + \phi_i(t)) \tag{4.1}$$

This payoff function shows that if a subject is not monitored, they receive $(300+\phi_i(t))$; if they are monitored they only receive $(300+\phi_i(t))$ if they meet the quota, which happens with probability $P_i(t,Q)$. In order to demonstrate some implications of this basic model I will assume simplifying assumptions for ϕ_i and P_i. Let $\phi_i(t) = a - bt$, and

$$P_i(t,Q) = \begin{cases} \frac{ct}{Q} & for \quad ct \leq Q \\ 1 & for \quad ct > Q \end{cases}$$

The results I derive are not special to these functional forms, but will hold for a broad range of functions satisfying the conditions: $\frac{d\phi_i}{dt} < 0$ and $\frac{dP_i}{dt} > 0$. Equation 4.1 now becomes:

$$\Pi_i(t,Q,m) = m\frac{ct}{Q}(300+a-bt) + (1-m)(300+a-bt) \tag{4.2}$$

The general problem is to maximize Π_i given some $t \leq T$.

$$\max_{t\leq T}\Pi_i(t,Q,m) = mP_i(t,Q)(300+\phi_i(t)) + (1-m)(300+\phi_i(t)) \tag{4.3}$$

The solution is:

$$t^* = \frac{1}{2}\frac{300mc + mca - bQ + mbQ}{mcb}. \tag{4.4}$$

Given this solution I can show the following points.

Implication 1: *When the monitoring level is zero, a subject should not work on Task A.*

When there is no monitoring, equation 4.1 becomes:

$$\Pi_i = (300+\phi_i(t)) \tag{4.5}$$

As stated earlier, the function $\phi_i(t)$ is declining everywhere in t, any time spent on Task A means less earnings can be made in Task B. So, $\frac{d\Pi_i}{dt} < 0 \forall t$. Therefore, to max $U_i = (300 + \phi_i(t))$, a subject should choose $t = 0$, and we should observe subjects playing only Task B.

Implication 2: *As the monitoring level increases, a subject should spend more time on Task A.*

$$\frac{dt^*}{dm} = \frac{d\left(\frac{1}{2}\frac{300mc+mca-bQ+mbQ}{mcb}\right)}{dm} = \frac{1}{2}\frac{Q}{m^2c} > 0 \tag{4.6}$$

As the monitoring level increases, shirking behavior is more likely to be punished. To avoid this punishment, a subject has to meet the quota. So, as the monitoring level increase, a subject should spend more time on Task A.

Implication 3: *If the subject chooses to work on Task A, once he meets the quota, he should spend no additional time on Task A.*

If $\exists\ \widehat{t}\ \ s.t.\ P_i\left(\widehat{t}, Q\right) = 1$, any $t^* > \widehat{t}$ will reduce the subjects expected utility. Let $P_i\left(\widehat{t}, Q\right) = 1$, and let $t^* > \widehat{t}$.

If subject chooses $t = \widehat{t} : \Pi_i = m\left(300 + \phi_i(\widehat{t})\right) + (1 - m)\left(300 + \phi_i(\widehat{t})\right) = 300 + \phi_i(\widehat{t})$.

If subject chooses $t = t^* : \Pi_i = m\left(300 + \phi_i(t^*)\right) + (1 - m)\left(300 + \phi_i(t^*)\right) = 300 + \phi_i(t^*)$.

$300 + \phi_i(\widehat{t}) > 300 + \phi_i(t^*)$ since $t^* > t'$ and ϕ_i is decreasing in t.

Implication 4: *For a given monitoring level, the subject will increase observed effort as quota increases until some point, as the quota gets relatively large, they will then exert less effort on Task A.*

Given that:

$$P_i(t, Q) = \begin{cases} \frac{ct}{Q} & for \quad ct \le Q \\ 1 & for \quad ct > Q \end{cases}$$

we have to consider that a subject should never choose

$$t^* = \frac{1}{2}\frac{300mc + mca - bQ + mbQ}{mcb} > \widehat{t}$$

where $\widehat{t}$ is defined by $P_i\left(\widehat{t}, Q\right) = 1$. In my example $P_i\left(t, Q\right) = 1$ when $ct = Q$. Therefore $\widehat{t} = \frac{Q}{c}$. The optimal amount of effort a subject will choose is

$$\min(t^*, \widehat{t}) = \min\left(\frac{1}{2}\frac{300mc + mca - bQ + mbQ}{mcb}, \frac{Q}{c}\right).$$

$$t^* < \widehat{t} \text{ when } Q > \frac{300 + a}{b\left(m+1\right)}mc$$

$$\frac{d\widehat{t}}{dQ} = \frac{d\left(\frac{Q}{c}\right)}{dQ} = \frac{1}{c} > 0 \tag{4.7}$$

$$\frac{dt^*}{dQ} = \frac{d\left(\frac{1}{2}\frac{300mc+mca-bQ+mbQ}{mcb}\right)}{dQ} = \frac{-1+m}{mc} < 0 \tag{4.8}$$

We see in Equation 4.7 that for $Q > \frac{300+a}{b(m+1)}mc$ a subject's effort will be increasing with Q, and in Equation 4.8 that for $Q < \frac{300+a}{b(m+1)}mc$ a subjects effort will be decreasing in Q. Fundamentally what this is saying is that subjects will maximize their probability of earning the 300 ECU wage when the quota is relatively low, but at some point increasing the quota will lead to a reduction in the effort (amount of numbers typed) given to Task A.

Implication 5: *If the subject cannot meet the quota even if they were to spend all of their time on Task A, he should not spend any time on Task A.*

If $P\left(t, Q\right) = 0 \;\; \forall t \leq T$, we see that equation 4.1 reduces to $U_i = \left(1 - m\right)\left(300 + \phi_i(t)\right)$. Again, the subject should only work on Task B.

Implication 6: *If the subject cannot meet the quota even if they were to spend all of their time on Task A, and the monitoring level is 100%, observations of the subjects playing either task would imply that they would rather engage the experiment than sit still.*

If $m = 1$, equation 4.1 becomes:

$$\Pi_i = P_i\left(t, Q\right)\left(300 + \phi_i(t)\right) \tag{4.9}$$

If $P(t, Q) = 0 \forall t \leq T$, then equation 4.9 is equal to zero. The subject gets zero financial gain from either task, so the subject should not engage either task unless they prefer it to sitting still.

Given these implications, we should expect the subjects to only play Task B when the monitoring level is zero, and to increase their effort on Task A as monitoring is increased. The subjects should also increase their effort on Task A when the quota increases if the quota is relatively low. An increase in the quota beyond some point will lead the subjects to reduce effort on Task A. If a subject meets the quota in a given period, he should then switch to playing only Task B as he no longer has any financial incentive to play Task A. When the quota is unattainable, the subjects should not try to reach the quota but instead spend all of their time on Task B. If the quota is unattainable and the subject is definitely going to be monitored, any engagement of either Task A or Task B implies that the subject prefers engaging the experiment to sitting still.

4.4 Experimental Results

Result 1: *When the monitoring level is zero, ninety-one percent of the subjects do not meet or exceed the quota. However, when the quota is small, many subjects meet that quota.*

Table 4.3 shows that subjects are highly unlikely to exert effort when they are not being monitored. Whereas *Implication 1* suggests that subjects should exert zero effort when the monitoring level is set to zero, the data show that some subjects do meet the quota when the monitoring level is 0%. Dickinson and Villeval (2005) found that 25% of subjects contributed at or above the desired output level; I find less than that. Table 4.3 shows that there are 14 out of 152 (09%) observations where the subject meets or exceeds the quota. But, we have to note that not all subjects could reach all quotas. Only one subject in the entire experiment met $Q = 25$, five subjects met $Q = 18$, seventeen subjects met $Q = 15$, all thirty two subjects met $Q = 8$, and $Q = 12$. So, to get another estimate of worker effort when monitoring level is zero I can use the the highest observed effort level for each subject to note if someone had the ability to meet a given quota. This leads to 14 out of 87 (16%)

Table 4.3: Subject effort when monitoring level is zero

number of four-digit numbers typed when m=0%								
Quota	0-5	6-10	11-15	16-20	21-25	NC=Q	NC>Q	Period
8	21	8	3	0	0	7	3	7
12	26	2	3	1	0	2	1	10
15	29	2	1	0	0	1	0	17
18	21	1	2	0	0	0	0	21
25	14	13	5	0	0	0	0	2

workers meeting or exceeding the quota when the monitoring level is zero. This estimate is not perfect as it has the ability to underestimate shirking. Some subjects may have been able to meet quotas that they were not observed to meet if they had worked harder. Best stated, when monitoring level is zero, the percentage of workers that meet or exceed the quota is in the range of 9% to 16%. Furthermore, most of the observations where workers meet or exceed the quota, 10 out of 14 (71%) , occur when the quota is relatively low (Q=8). When $Q > 8$ and monitoring is zero, only 4 out of 55 (07%) subjects meet or exceed the quota.

To test *Implications 2 and 4,* I run a linear random effects panel regression. This type of regression controls for omitted variables that differ between subjects and omitted variables that vary within subjects over time. I will regress the amount of four-digit numbers typed on monitoring level (*Implication 2*), quota less than 15, and quota greater than 15 (*Implication 4).* The regressor, $dummy_{Q<15}$, is set to 1 when $Q < 15$ and zero when $Q > 15$, $dummy_{Q>15}$ is set to 1 when $Q > 15$ and zero when $Q < 15$. If subjects respond in the experiment as predicted by *Implication 4* they will increase effort up until some Q, and then begin to reduce effort. If this is the case, I need a constant for both situations. The regressor $dummy_{period \leq 5}$ is set to 1 for periods 1-5 and zero otherwise. This is used to determine if behavior in the first five periods is significantly different than play in latter periods. This could be a factor if the subjects experience any learning effects in the first five periods. Prior to the experiment, it is unlikely that the subjects know how many four-digit-numbers they can type in 45 seconds; the subjects may use the initial periods to gauge their ability. Table 4.4 reports the results

Table 4.4: Results of linear random effects panel regression of number of four-digit numbers typed on possible explanatory variables.

	Coef.	**Std. Err**	**P-value**
Monitoring	.0786	.0040	0.000
quota(<15)	.5140	.0742	0.000
quota(>15)	-.9551	.0440	0.000
$dummy_{Q<15}$	1.404	.8858	0.281
$dummy_{Q>15}$	6.257	1.002	0.000
$dummy_{period \leq 5}$	1.850	.3807	0.000
Number of Groups	32		
Observations/Group	30		

of this regression.

Result 2: *Monitoring has a positive significant impact on effort given to task 1.*

Table 4.4 shows that as monitoring increases subjects choose to exert more effort (type more numbers) on the monitored task. This result is consistent with *Implication 2.* Specifically, if monitoring goes up by one percentage point, the average subject increases the amount of four-digit numbers by 7.9%. Figure 4.1 is a scatterplot of four-digit-numbers typed on quota for a given monitoring level. Figure 4.1 allows us to see the impact of monitoring in a sequence. There are 960 observations in each cell. The data have been jittered due to the high level of overlap of the data points. The dark spots in each cell represent the most overlap. We see that when the monitoring level is low, observations are clustered around zero effort. As monitoring increases the mass of observations starts to move up, and when monitoring is at 100% we see a very strong correlation between the quota and the amount typed. If we look at Figure 4.1 we see that the subjects are exerting effort where monitoring is zero and $Q = 25$. The subjects saw this combination of quota and monitoring level early in the experiment (period 2). The significance of the regressor $dummy_{period \leq 5}$ in Table 4.4 shows that subjects were playing differently early in the experiment. It is possible that the subjects were exploring the game space.

Result 3: *Once subjects meet the quota in Task A, most switch to Task B exclusively.*

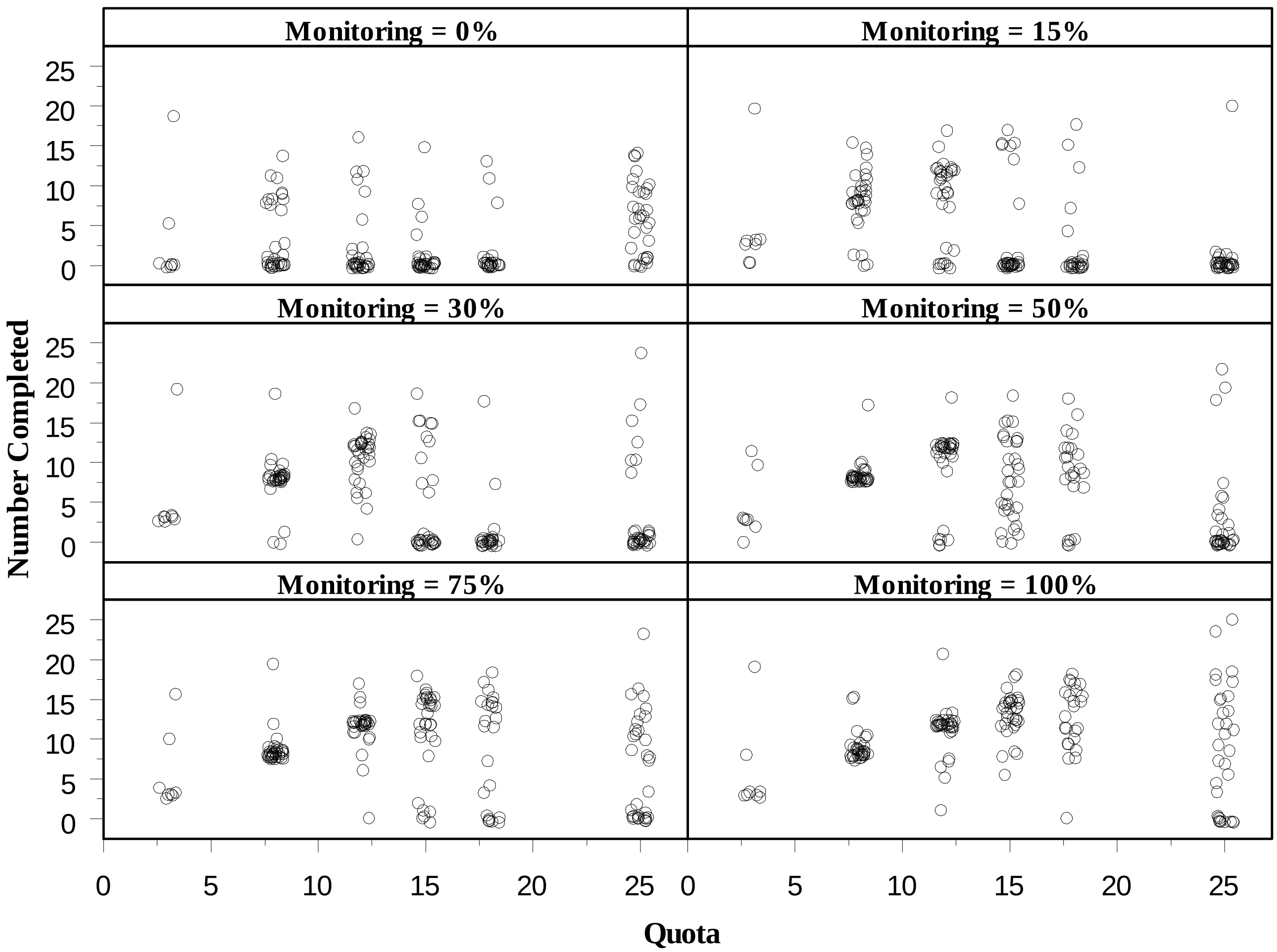

Figure 4.1: Subject effort given a constant monitoring level

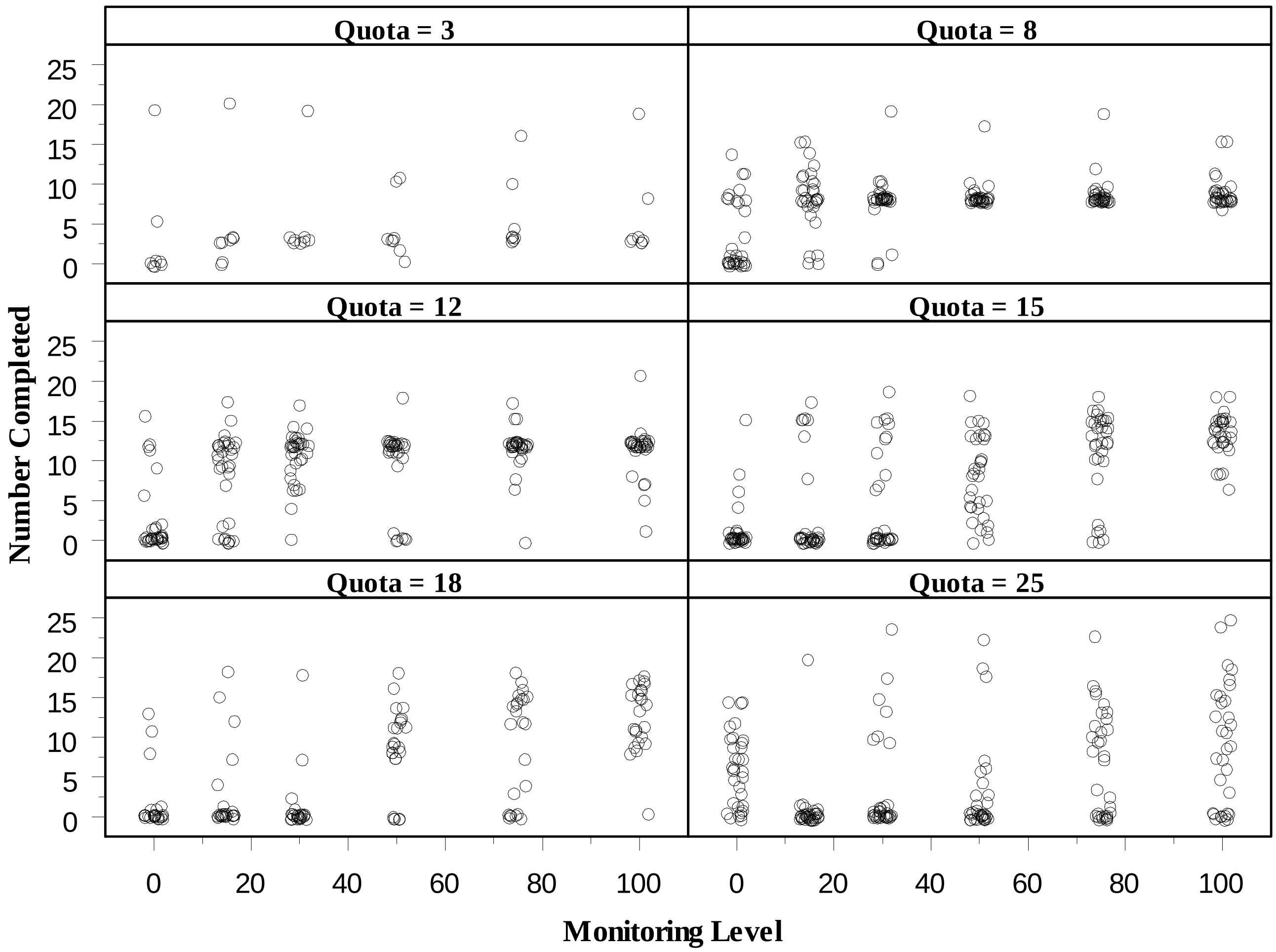

Figure 4.2: Subject effort given a constant quota

If we refer back to Figure 4.2 we see few observations where the amount typed is greater than the quota. Consider the case where monitoring is equal to 100%, there are 70 instances where the amount typed is greater than or equal to the quota, of these 70 instances, 62 hit the quota exactly. In Figure 4.2, we can observe that the number typed is rarely above the quota.

Result 4: *The quota level, for quota less than 15, has a positive significant impact on effort. The quota level, for quota greater than 15, has a negative significant impact.*

Table 4.4 shows that, consistent with *Implication 4,* when the quota is relatively small subjects increase their effort on Task A. However, when the quota is greater than 15 subjects' effort is decreasing as the quota increases. We can see in Figure 4.2 that when the quota increases from 8 to 12, there is an increase in the amount of effort given to Task A. When the quota increases to 15, we can see a drop off in the effort given to Task A, particularly when monitoring is less than 50%.

Result 5: *When the quota is set to 25, this is effectively* $P(t,Q) = 0\ \forall t$*, and monitoring is less than 100%, subjects prefer to spend more time on Task B than Task A.*

Prior to the experiment, the quota level of 25 was believed to be so high that no subject could hit it, and would therefore be implicitly unattainable. This held true for 30 of the 32 subjects, but one subject was able to reach $Q = 25$ and another was able to type 24 four-digit numbers. In the minds of these two subjects, the quota of 25, was probably not unattainable. The other subjects never typed as much as 20 four-digit numbers. So, for 30 of the 32 subjects, $Q = 25$ should have been inferred as unattainable. *Implication 3* demonstrated that if a subject could not meet the quota even if they spent all of their time on Task A, then they should not try to meet the quota. Table 4.5 shows that when subjects think they cannot meet the quota, they will not try to meet the quota. If we contrast this unattainable quota with the high quota of 18, we see that the subjects exert much more effort when there is at least some possibility that they might reach the quota. Consider, for example, the case where monitoring = 50%.When the quota is 25, the number of subjects that type five or fewer four-digit numbers is 26 out of 32 (81%); when the quota is 18, the number of subjects that type five or fewer four-digit numbers is 5 out of 24 (21%). If we

Table 4.5: Subject effort when Q=25

number of four-digit numbers typed when Q=25							
Mon. Level	0-5	6-10	11-15	16-20	21-25	NC=Q	Period
0%	14	13	5	0	0	0	2
15%	31	0	0	1	0	0	24
30%	25	3	2	1	1	0	13
50%	26	3	0	2	1	0	19
75%	15	6	8	2	1	0	8
100%	12	5	9	4	2	1	30

Table 4.6: Subject effort when Q=18

number of four-digit numbers typed when Q=18[4]							
Mon. Level	0-5	6-10	11-15	16-20	21-25	NC=Q	Period
0%	21	1	2	0	0	0	21
15%	20	1	2	1	0	1	20
30%	22	1	6	1	0	1	28
50%	5	9	8	2	0	1	6
75%	9	1	11	3	0	1	26
100%	1	6	10	7	0	1	23

look to Figure 4.2 we can see that effort given to Task A is much greater when the quota is 15 compared to when the quota is 25. *Result 1* showed that some subjects will meet easily attainable quotas even without financial incentives. *Result 5* shows that unattainable quotas will cause subjects to greatly reduce effort.

Result 6: *When the Quota is set to 25 ($P(t,Q) = 0 \quad \forall t$), and the monitoring level is set to 100%, subjects do not sit still, they engage both Task A and Task B.*

Table 4.7 shows when monitoring is 100% and $Q = 25$ many subjects exert effort even though they know they will not meet the quota. One subject was extremely fast and did meet the quota, another made it to 24. No one else made it to 20. Seven subjects engaged neither task, two subjects engaged both, three subjects played only the outside option, and twenty engaged only the typing task. Similar to this result, both Cadsby et al. (2004), and Dickinson and Villeval (2005) found subjects contributing without financial incentives.

Table 4.7: Subject effort when monitoring level is 1

number of four-digit numbers typed when m=100%								
Quota	0-5	6-10	11-15	16-20	21-25	NC=Q	NC>Q	Period
8	0	28	4	0	0	27	4	5
12	2	3	26	0	1	24	1	22
15	0	4	25	3	0	9	3	14
18	1	6	10	7	0	1	0	23
25	12	5	9	4	2	1	0	30

Result 6 lends support to the claim that subjects will engage tasks simply because they are bored. One could argue that since the only possible way a subject could get any earnings was to try and reach the quota, effort given to Task A was not play due to boredom. While this is plausible, it is not likely given the intensity of effort given to Task A. If we look at Tables 4.5 and 4.6, when monitoring is set to 100%, we see that work intensity is higher for $Q = 18$. Consider that when $Q = 25$ only 15 out of 32 (47%) subjects type more than 10 numbers, but when $Q = 18$ 17 out of 24 (71%) subjects type more than 10 numbers. So, the workers who choose to give effort to Task A when monitoring is 100% and $Q = 25$ do not appear to be exerting high effort, and that is consistent with play due to boredom.

4.5 Conclusion

This paper presented an experimental study of worker productivity. The main objective of the paper was to examine the effect of various incentive schemes on subject behavior in an environment that is meant to mimic a work setting where the worker has the ability to shirk by engaging in some task other than work for the employer. A secondary objective of the paper was to examine past experimental results that showed subjects exerting effort when no financial incentive to do so existed, and determine if this was consistent with play due to boredom.

The key finding of this paper is that workers in a laboratory work setting with an outside option available to them do shirk, but monitoring is quite successful at reducing shirking.[5] In fact, when the quota is not difficult to attain, very little monitoring is necessary to gain subject compliance. However, when the quota is unattainable, the subjects revolt and exert very little effort. When there is no monitoring, I do observe that some subjects may still have some intrinsic motivation to play the primary task, but less than previous work reported. Another finding is that by allowing subjects to participate in an outside option, I am able to mitigate play that is due to boredom. And by implicitly disallowing earnings we are able to see that subjects prefer to play tasks instead of sitting still.

Subjects in this experiment were observed to work until they met their quota and then switched to the outside option. This paper began with the story of a worker who was fired for a lack of productivity. But it is not clear that he was shirking. Once he completed his assignments, or met his quota, he played the outside option available to him. This is the exact behavior subjects exhibited in the laboratory.

Future work will examine if subjects who are monitored over any shirking behavior exert more effort on their primary task than subjects monitored over performance goals. The issue of self selection will also be examined by paying subjects based on the amount of numbers they type as opposed to a flat wage. Subjects will then be allowed to choose the pay for performance scheme or the flat wage and monitoring scheme.

4.6 Appendix: Experiment Instructions

In this part of the experiment you will be able to work on two tasks. You can split your time among the tasks however you choose. Meaning you can spend all of your time on task A and none on task B, all of your time on task B and none on task A, or some combination of task A and task B. Each round in this experiment is 45 seconds long.

In the box marked task A on your handout a number is displayed. Your task is to type the number you see in the box provided. Every time you click the "OK" button, a new number will come up and you can go through the task again. In each round of the experiment, the

[5] This is also consistent with a model by Stevens and Thevaranjan (2007) that allows for ethical behavior.

default payment for task A is 300 ECUs (1000 ECUs = \$1.00). Payment will be further described below.

If you look at your handout, you will see two other pieces of information for task A (Quota and Probability of being monitored).

Monitoring means that the computer is checking to make sure that you have met the quota (defined in next paragraph). The "monitoring level" tells you how likely it is that the computer is going to check on you. If the "monitoring level" is 0%, that means the computer will not check on you. If it is 100%, that means the computer will definitely check on you.

If you are monitored: The "quota" is the minimum number of task A numbers you have to type in order to earn 300 ECUs. In other words, to earn the 300 ECUs, you must type at least the amount of numbers specified by the quota and they must be typed correctly. Your screen will update the amount of completed correct numbers you have typed ("number completed" on your handout). If you do not type enough numbers correctly to cover the quota, your earnings for the round are zero.

If you are not monitored: If you meet the quota you earn 300 ECUs. If you don't meet the quota, you earn 300 ECUs.

The monitoring level is stated at the beginning of every round. At the end of each round, the computer will generate a random number between 1 and 100 with all numbers being equally likely to determine whether or not you are in fact monitored. For example, assume you enter the round and the "monitoring level" is set to 40 (this is a 40% chance that you will be monitored). If the number drawn at the end of the round is between 41 and 100, you will not be monitored. If the number drawn is between 1 and 40, you will be monitored.

Alternatively, you may also choose to work on task B (the bottom section of the handout).

In this task you will be playing a game against the computer. The computer is programmed to pick either option H or option T. You also have the ability to pick option H or option T (you do this by choosing H or T and then clicking the "ok" button).

If both of you pick the same option you win, and earn 2 ECUs, if the choices don't match, the computer wins, and you earn zero ECUs. However, if you are monitored in task A, and you did not meet the quota, whatever earnings you had accumulated in task B will be wiped out.

The example in the handout has the monitoring level at 25 (that is a 25% chance that you will be monitored), and the quota is set at 2. If the "number completed" at the end of the 45 second round is greater than the quota, or equal to the quota, you will earn 300 ECUs from task A plus whatever you earned on task B. If "number completed" is less than the quota, two things can happen:

1. You are monitored, and you earn zero for the round (zero from task A and zero from task B)

2. You are not monitored, and you earn 300 ECUs from task A plus whatever you earned on task B.

On the second page of your handout, the example shows a case where monitoring occurred. However, since the quota was met, the earnings from task A are 300 ECUs. Your total earnings are 306, because you earned 6 ECUs from task B. If the quota had not been met, the task A earnings would be zero, and your earnings from task B would be wiped out, so your total earnings for the round would be zero.

CHAPTER 5

Concluding Remarks

This dissertation had three primary objectives: examining the consistency of preferenes across instiutions; investigating shirking behavior in a real effort laboratory experiment; and examing if past experimental results that are at odds with theory could be artifacts of experimental design.

The key contribution of chapter 3 is that I am able to show that subjects have consistent risk preferences across institutions. This is consistent with standard assumptions, but past experiments had called these assumptions into question. I first show that my subject pool is similar to past subject pools, and then bu using both non-parametricall and parametrical tests that my subjects have consistent risk preferences. This result lends support to the argument that subjects' bids in first price sealed bid auctions are due to risk aversion, and not simply consistent with a risk aversion model.

Future work will need to focus on why past experiments did not find consistency. Specifics of these institutions along with the experimental design (in terms of both complexity of the instructions, and complexity of the institution) will need to be examined. This will be important for policy decisons that attempt to take experimental and theoretical results into the field. It could be the case that individuals are consistent if the they understand the rules, so any implementation of policy should strive for clarity.

The key contribution of chapter 4 is that workers do shirk, but monitoring and an attainable quota work well at mitigating shirking behavior. However, when the quota is difficult to reach, subjects are more likely to shirk. So managers have to walk a fine line in the intensity of work they require. It is possible that by requiring less work they end up with more productivity. I also find evidence that past experimental results were artifacts of their experimental design. I designed parts of my experiment to test if workers prefer to engage a task as opposed to sitting still, and I observe that subjects do, in fact, prefer to engage the experiment. In any future experiments, the experimental design should take pains to ensure that 'costly' participation overrides the cost of sitting still.

Future work in this area will examine different types of monitoring and payment schemes. Specifically, I will examine the effects of monitoring on any shirking behavior. Managers may feel that workers should not shirk at all, and any engagement of the outside options would be punished. So instead of monitoring over some quota, they just check to make sure that you are *always* working on the main task. It is possible that this could lead workers to rebel in a more passive way. What we saw in my results was that with an outside option and an attainable quota, subjects would work intensly to meet their quota and then play the outside option. It could be the case that monitoring both tasks leads to less shirking in terms subjects engaging the outside option, but this zealous monitoring could lower the intensity of work in the primary task.

I will also examine a piece rate wage scheme. I will do this by allowing subjects to self select into a pay for performance scheme or the current flat wage scheme. This may increase productivity because those that are best at the primary task will self select into it. This is similar to stock brokers, who work on commission, and securities analysts, who work on salary, but both are employed by the same firm. Once I have results on all three of these sytems in terms of productivity, I can them analyze them from the cost side. It may not be the case that the most productive scheme in terms of output leads to the highest profits for the firm.

APPENDIX

Human Subjects Committee Approval

Florida State
UNIVERSITY

Office of the Vice President For Research
Human Subjects Committee
Tallahassee, Florida 32306-2742
(850) 644-8673 · FAX (850) 644-4392

APPROVAL MEMORANDUM

Date: 5/10/2006

To:
Russell Engel
Dept. of Economics, Bellamy Bldg.

Dept.: **ECONOMICS**

From: **Thomas L. Jacobson, Chair**

Re: **Use of Human Subjects in Research**
The Impact of Monitoring and Quotas on Subjects' Effort Level

The forms that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Secretary, the Chair, and two members of the Human Subjects Committee. Your project is determined to be Exempt per 45 CFR § 46.101(b) 2 and has been approved by an accelerated review process.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals, which may be required.

If the project has not been completed by **5/9/2007** you must request renewed approval for continuation of the project.

You are advised that any change in protocol in this project must be approved by resubmission of the project to the Committee for approval. Also, the principal investigator must promptly report, in writing, any unexpected problems causing risks to research subjects or others.

By copy of this memorandum, the chairman of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols of such investigations as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Protection from Research Risks. The Assurance Number is IRB00000446.

Cc: Mark Isaac
HSC No. 2006.0358

Florida State
UNIVERSITY

Office of the Vice President For Research
Human Subjects Committee
Tallahassee, Florida 32306-2742
(850) 644-8673 · FAX (850) 644-4392

APPROVAL MEMORANDUM

Date: 4/20/2006

To:
Edward Cokely Russell Engel
MC 2180

Dept.: **ECONOMICS**

From: **Thomas L. Jacobson, Chair**

Re: **Use of Human Subjects in Research**
A test of correlation between risk aversion and ellicitation devices

The forms that you submitted to this office in regard to the use of human subjects in the proposal referenced above have been reviewed by the Secretary, the Chair, and two members of the Human Subjects Committee. Your project is determined to be Exempt per 45 CFR § 46.101(b) 2 and has been approved by an accelerated review process.

The Human Subjects Committee has not evaluated your proposal for scientific merit, except to weigh the risk to the human participants and the aspects of the proposal related to potential risk and benefit. This approval does not replace any departmental or other approvals, which may be required.

If the project has not been completed by **4/16/2007** you must request renewed approval for continuation of the project.

You are advised that any change in protocol in this project must be approved by resubmission of the project to the Committee for approval. Also, the principal investigator must promptly report, in writing any unexpected problems causing risks to research subjects or others.

By copy of this memorandum, the chairman of your department and/or your major professor is reminded that he/she is responsible for being informed concerning research projects involving human subjects in the department, and should review protocols of such investigations as often as needed to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

This institution has an Assurance on file with the Office for Protection from Research Risks. The Assurance Number is IRB00000446.

Cc: Mark Isaac, Anders Ericsson
HSC No. 2006.0215

REFERENCES

[1] Akerlof, G.A. 1982. Labor contracts as partial gift exchange. Quarterly Journal of Economics 97: 543-569.

[2] Becker, Gordon M., Morris H. Degroot, and Jacob Marschak. (1964). "Measuring Utility by a Single-Response Sequential Method," Behavioral Science 9, 226-232.

[3] Berg Joyce & Dickhaut John & McCabe Kevin, 1995. "Trust, Reciprocity, and Social History," Games and Economic Behavior, Elsevier, vol. 10(1), pages 122-142, July.

[4] Beveridge, W. I. B. 1950. The Art of Scientific Investigation. New York: Random House (Vintage Books).

[5] Cadsby et al. (2004). The Effects of Compensation Schemes on Self-Selection and Work Productivity: An Experimental Investigation. *Experiments in Economic Sciences - New Approaches to Solving Real-world Problems* (Working Paper)

[6] Camerer, Colin, "Individual Decision Making," in J.H. Kagel and A.E. Roth (eds.), The Handbook of Experimental Economics, Princeton: Princeton University Press (1995).

[7] Cason Timothy N., 1993. "Seller Incentive Properties of EPA's Emission Trading Auction," Journal of Environmental Economics and Management, Elsevier, vol. 25(2), pages 177-195, September.

[8] Charness, Gary & Frechette, Guillaume R. & Kagel, John 2004. "How Robust is Laboratory Gift Exchange?," Experimental Economics, Springer, vol. 7(2), pages 189-205, 06.

[9] Cox, James C., Bruce Roberson, and Vernon L. Smith. (1982). "Theory and Behavior of Single-Object Auctions," In V. L. Smith (ed.), Research in Experimental Economics, Vol. 2. Greenwich, Conn.: JAIPress, Inc.

[10] Cox, James C., Vernon L. Smith, and James M. Walker. (1988). "Theory and Individual Behavior of First-Price Auctions," Journal of Risk and Uncertainty 1, 61-99.

[11] Cox, James C & Smith, Vernon L & Walker, James M, 1992. "Theory and Misbehavior of First-Price Auctions: Comment," American Economic Review, American Economic Association, vol. 82(5), pages 1392-412, December.

[12] Davis, D. D & Holt, C.A. 1993. Experimental Economics. Princeton: Princeton University Press.

[13] Dickinson, David L. (1999) "An Experimental Examination of Labor Supply and Work Intensities." Journal of Labor Economics, 17(4), part 1: 638-670.

[14] Dickinson, D. and M. Villeval (2005). Does Monitoring Decrease Work Effort? The Complementarity Between Agency and Crowding-Out Theories,Working Papers 05-12, Department of Economics, Appalachian State University.

[15] Engelbrecht-Wiggans, R., "The Effect of Regret on Optimal Bidding in Auctions," Management Science, 1989, Vol. 35(6), pp. 685-692.

[16] Engelmann D and A Ortmann (2002), The Robustness of Laboratory Gift Exchange. Manuscript.

[17] Falk, A., Ichino, A., 2006. Clean Evidence on Peer Effects. Journal of Labor Economics 24, 39-57.

[18] Fehr, E., G. Kirchsteiger, and A. Riedl. 1993. Does fairness prevent market clearing? An experimental investigation. The Quarterly Journal of Economics 108: 437-459.

[19] Fehr, E., E. Kirchler, A. Weichbold, and S. Gachter. 1998. When social norms overpower competition: Gift exchange in experimental labor markets. Journal of Labor Economics 16: 324-351.

[20] Fischbacher, Urs (1999). z-Tree - Zurich Toolbox for Readymade Economic Experiments - Experimenter's Manual, Working Paper Nr. 21, Institute for Empirical Research in Economics, University of Zurich.

[21] Friedman, Daniel, 1992. "Theory and Misbehavior of First-Price Auctions: Comment," American Economic Review, American Economic Association, vol. 82(5), pages 1374-78, December.

[22] Gneezy, Uri, Niederle M. and Rustichini A., Performance in competitive environments: gender differences, Quarterly Journal of Economics 118 (2003), pp. 1049–1074.

[23] Goeree, Jacob K., Charles A. Holt, and Thomas Palfrey (1999) "Quantal Response Equilibrium and Overbidding in Private-Value Auctions," Discussion Paper, California Institute of Technology.

[24] Grossman, S. J., and O.D.Hart (1983). An Analysis of the Principal-Agent Problem. *Econometrica* 51(1), pages 7-45

[25] Guillen, Pablo and Veszteg,Robert F. 2006. "Subject Pool Bias in Economics Experiments," ThE Papers 06/03, Department of Economic Theory and Economic History of the University of Granada.

[26] Harrison, Glenn W, 1989. "Theory and Misbehavior of First-Price Auctions," American Economic Review, American Economic Association, vol. 79(4), pages 749-62, September.

[27] Harrison, Glenn W. (1990). "Risk Attitudes in First-Price Auction Experiments: A Bayesian Analysis," Review of Economics and Statistics 72, 541-546.

[28] Hey, John D., 1991. "Advances in behavioral economics: Leonard Green and John H. Kagel (eds.), Vol. 2. Ablex, Norwood, IL, 1990. pp. xiv + 325," Journal of Economic Psychology, Elsevier, vol. 12(4), pages 725-726, December.

[29] Holt, Charles. 1980. "Competitive Bidding for Contracts under Alternative Auction Procedures," Journal of Political Economy 88, pp. 433-445.

[30] Holt, C. and S. Laury (2002). "Risk Aversion and Incentive Effects," American Economic Review 92 (5), 1644-1655.

[31] Isaac, R. M. and D. James (2000). "Just Who Are You Calling Risk Averse?," Journal of Risk and Uncertainty 20 (2), 177-187.

[32] Ivanova-Stenzel, Radosveta and Timothy C. Salmon. (2005) Revenue Equivalence Revisited. Working Paper Florida State University.

[33] Kagel, John H.; Harstad, Ronald M. and Levin, Dan. (1987) "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study." Econometrica, November 1987, 55(6), pp. 1275-304.

[34] Kagel, John and Levin, Dan (1993): "Independent Private Value Auctions: Bidder Behavior in First-, Second- and Third Price Auctions with Varying Numbers of Bidders," Economic Journal 103, 868-879.

[35] Kagel, John, "Auctions: A Survey of Experimental Research," in J.H. Kagel and A.E. Roth (eds.), The Handbook of Experimental Economics, Princeton: Princeton University Press (1995).

[36] Knight, F.H.(1921) Risk, Uncertainty and Profit. Boston, New York: Houghton Mifflin.

[37] Lei, V., C. Noussair, and C. Plott (2001). Non-Speculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality *Econometrica* 69(4), pp. 831-859.

[38] McAfee, R Preston & McMillan, John, 1987. "Auctions and Bidding," Journal of Economic Literature, American Economic Association, vol. 25(2), pages 699-738, June.

[39] Montmarquette, Claude & Rullière, Jean-Louis & Villeval, Marie-Claire & Zeiliger, Romain, 2004. "Redesigning Teams and Incentives in a Merger: An Experiment with Managers and Students," IZA Discussion Papers 1057, Institute for the Study of Labor (IZA).

[40] Milgrom, Paul and Robert Weber, "A Theory of Auctions and Competitive Bidding," Econometrica (1982), 50: pp. 1089-1122.

[41] Mirrlees, J. A.(1975). The Theory of Moral Hazard and Unobservable Behaviour I, Mimeo, Oxford (Published in review of Economic Studies in 1999).

[42] Myerson, Roger. "Optimal Auction Design." Mathematics of Operations Research. February 1981, 6 (1), pp. 58-73.

[43] Nagin, D. et al. (2002). Monitoring, Motivation and Management: The Determinants of Opportunistic Behavior in a Field Experiment. *American Economic Review* 92 (4), pp. 850-73.

[44] Nalbantian, H.R., and A. Schotter(1997). Productivity under Group Incentives: An Experimental Study. *American Economic Review* 87, 314-341.

[45] Palfrey, T. and S. Pevnitskaya (2003). Endogenous Entry and Self-selection in Private Value Auctions: An Experimental Study - Caltech Social Science Working Paper No.1172, 2003.

[46] Plott, Charles (1982). "Industrial Organization Theory and Experimental Economics." Journal of Economic Literature, 20: 1485–1527.

[47] Prasad, K and T. C. Salmon (2007). "Self Selection and Market Power in Risk Sharing Contracts," Working Paper Florida State University.

[48] Riley, John and William Samuelson. 1981. "Optimal Auctions," American Economic Review 71, pp. 381-392.

[49] Roth, A.E., "Laboratory Experimentation in Economics, and Its Relation to Economic Theory," Scientific Inquiry in Philosophical Perspective, Nicholas Rescher, editor, Lanham, University Press of America, 1987, 147-167.

[50] Samuelson, Larry 2005. "Economic Theory and Experimental Economics," Journal of Economic Literature, American Economic Association, vol. 43(1), pages 65-107, March.

[51] Sillamaa, M. A., 1999. "How work effort responds to wage taxation: An experimental test of a zero top marginal tax rate," Journal of Public Economics, Elsevier, vol. 73(1), pages 125-134, July.

[52] Smith V.L. (1962), "An experimental study of competitive market behavior", Journal of Political Economy 70, 111-137.

[53] Smith V.L. (1982), "Microeconomic systems as an experimental science", American Economic Review 72, 923-955.

[54] Smith, Vernon, Gerry Suchanek, and Arlington Williams, 1988, Bubbles, crashes, and endogenous expectations in experimental spot asset markets, Econometrica 56, 1119-1151.

[55] Smith, Vernon L. and James M. Walker (1993) "Rewards, Experience, and Decision Costs in First-Price Auctions," Economic Inquiry, 31, 237-245.

[56] Stevens, D. and Thevaranjan, A. (2007) FSU Working Paper.

[57] Taleb, Nassim, The Black Swan: The Impact of the Highly Improbable. New York : Random House, 2007.

[58] van Dijk, F., Joep Sonnemans and Frans van Winden (2001). Incentive Systems in a Real Effort Experiment. *European Economic Review* 45, 187-214.

[59] Vickrey, William. 1961. "Counterspeculation and Competitive Sealed Tenders." Journal of Finance. 16:1, pp. 8-37.

[60] Walker, James M., Vernon L. Smith and James C. Cox. (1987). "Bidding Behavior in First Price Sealed Bid Auctions: Use of Computerized Nash Competitors," Economics Letters 23, 239-244.

[61] Wolfstetter, E. (1995). "Auctions: An Introduction." Journal of Economic Surveys 10(4): 367-420.

BIOGRAPHICAL SKETCH

Russell P. Engel

Russell Paul Engel was born in Fall River, Massachusetts on February 14, 1979 to Cheryl Engel and William Franco III. He was raised primarily by his maternal grandparents Jimmy and Carol Engel, and also spent time with his paternal grandparents William and Helena Franco. Russell graduated from Durfee High School in June 1997. He then moved to Tallahassee, Florida to study economics at Florida State University. He earned his Bachelor of Science Degree in Economics in April 2001. Russell continued at FSU and earned a Master of Science Degree in August 2003, and the Doctor of Philosophy Degree in December 2007. He is now employed as an Assistant Professor of Economics in the John F. Welch College of Business at Sacred Heart University in Fairfield, Connecticut, and lives with his wife Judy Thuy Nguyen in Milford, Connecticut.