A STUDY OF HUMAN DECISION-MAKING IN ECONOMIC GAMES

Ellen P. Green

Dissertation submitted to the faculty of Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of:

Doctor of Philosophy
In
Economics, Science

Sheryl B. Ball
Suqin Ge
Joao C. Macieira
T. Nicolaus Tideman
Wen You

November 4th, 2011
Blacksburg, Virginia

A STUDY OF HUMAN DECISION-MAKING IN ECONOMIC GAMES

Ellen P. Green

ABSTRACT:

This dissertation contains three essays on the impact of other-regarding behavior on human decision-making. Chapter II uses experimental methods to analyze the relative performance of a variety of compensation contracts. This study creates an environment in which individuals are paid via common payment mechanisms employed in the dual-principal agent relationships (Piece Rate, Flat Rate, Salary, Bonus and Socialization) and examines the effect that different incentive structures have on agent behavior. In Chapter III I explore the potential outcomes of blended payment structures in a dual-principal agent environment. I draw from the previously conducted experimental study in Chapter II and simulate agent behavior induced by blended payment mechanisms. In Chapter IV, I move away from studying payment mechanisms to investigate the impact of intentionality and responsibility on an individual's decision-making process. I explore the effects of direct and indirect responsibility as well as selfish and kind intentions using experimental methodology. Each of these essays provides further evidence that other-regarding behavior has a significant impact on the outcome of an economic situation; therefore, emphasizing the need to address such behavior in theoretical designs.

DEDICATION:

I dedicate this dissertation to my family.

To the memory of my mother,
who has taught me more than I could ever attempt to reflect in the limited space here.

To my father,
who instilled in me the desire to strive for excellence in myself and not for anybody else.

To my best friend and husband,
without all of your unwavering encouragement this dissertation would not be possible.

And to my support team,
Ben and Sam,
you are the best brothers a sister could ask for.

ACKNOWLEDGEMENTS:

When they say it takes a village to raise a child, they should also say it takes a village to finish a dissertation. I do not believe that this dissertation would have been possible with out all of the immense support from my committee members.

First, I would like to express my gratitude to my advisor, Dr. Sheryl Ball, for her help in the guidance of my dissertation. Her knowledge of economics motivated and inspired me to complete my dissertation. I would like to thank Nic Tideman for his support both academically and personally during my time at Virginia Tech. I would also like to thank Joao Maciera, Sue Ge, and Wendy You, for their continued support of my dissertation and my education as a graduate student at Virginia Tech.

I would also like to thank Ben Congleton and Russell Osborne for their help in creating the computer interface necessary to conduct the experiments in Chapter II. I would like to thank Ryan Donnelly, for creating the essays for the Dual-Principal Agent Game.

In addition I would like to thank Landis Atkinson and Aida Crus-Folgar for all of their help running experiments throughout the Spring of 2011. I would also like to acknowledge Landis for learning to program in z-tree to develop the code for the experiment in Chapter IV.

TABLE OF CONTENTS:

LIST OF FIGURES:

LIST OF TABLES:

CHAPTER I: INTRODUCTION

ELLEN P GREEN
VIRGINIA TECH
DEPARTMENT OF ECONOMICS

The goal of this dissertation is to expand on the current economic literature that discusses the impact of other-regarding preferences on decision-making processes. The incorporation of other-regarding preferences in theoretical work has significantly progressed the accuracy of economic predictions in the past several decades. The motivation behind incorporating other-regarding preferences into theory work is the result of many economic experiments. The first experiments testing economic behavior in decision-making games started in the 1970s. Stahl (1972) developed the dictator game and showed that individuals do not behave in the purely selfish manner that is predicted by the self-regarding theory. This result inspired the exploration of the influence of other-regarding behavior continued through the years that confirmed the results of Stahl. A substantial literature in the decades following Stahl's experiment focused on defining the motivating factor behind the unselfish outcomes in many economic situations (i.e. altruism, fairness, inequity aversion).

As experimental literature grew and the methodology started to become more widely accepted by economist, investigations extended beyond the other-regarding behavior on simple decision-making tasks and began to explore other popular topics in economics, such as labor theory. In this dissertation, I explore the impact other-regarding behavior has on two economic scenarios; the first of which is in a dual-principal agent relationship.

In Chapters II and III, I investigate the influence of other-regarding preferences in a dual-principal agent relationship and the impact it has on the outcome of contract design. In particular, I investigate the efficiency of various payment mechanisms in the dual-principal agent relationship. This paper expands on the findings of earlier experimental economists, who focus on the typical principal-agent relationship and the sole effect of other-regarding behavior on contract design. In the dual-principal agent problem an agent has responsibilities to two different principals. For instance, a physician is an agent of their patients and an agent of their health care employer. A critical feature of the relationship that I investigate here is that there is an interior solution for the number of services provided to the client (the patient). While the agent is assumed to know this critical value, the employer only knows what the average value should be over all downstream principals, and the downstream principal, or client, has very little information.

In Chapter II, I use experimental methods to analyze the relative performance of a variety of compensation contracts. This study creates an environment in which individuals are paid via common payment mechanisms employed in the dual-principal agent relationships (Piece Rate, Flat Rate, Salary, Bonus and Socialization). I then examine the effect that different incentive structures have on agent behavior in the dual-principal agent relationship. In addition, I specifically address the effect that other-regarding behavior has on decision making in the physician-patient relationship. Results suggest that the existence of other-regarding behavior substantially affects the choices made by agents. The results also show that particular compensation contracts outperformed others on a variety of measures. From my experiments, we can show that classic contract theories are misleading in their conclusions. In this study, the payment mechanism that resulted in the highest quality service was the Salary contract; where in classic theory a salaried agent would underprovide services relative to the other compensation contracts. Standard contract theory emphasizes the need for a direct link between the final product provided by the agent and their compensation for services to encourage effort. Payment mechanisms without this link are predicted to result in inefficient agent effort and shirking. However, this study suggests that this link may not be necessary to achieve a high quality of service in the dual-principal agent problem. Because of the significant effect that other-regarding behavior has on an individual's utility, agents do not need to be directly incentivized to provide high quality services.

In Chapter III, to expand on the findings presented in Chapter II by using counterfactual simulations to demonstrate agent behavior induced by blended payment mechanisms. In the laboratory experiment, as discussed above, the effect of the Flat Rate, Piece Rate, Salary, Piece Rate with Bonus and Flat Rate with Bonus payment mechanisms on agent behavior was explored. By employing a set of axioms and modeling agent behavior, I am was able to show the optimal combinations of the aforementioned payment structures, lies in the combination of Flat Rate and Piece Rate payment mechanisms, in terms of employer profit.

Furthermore, the context and framing of an event can significantly impact the decision-making process. This topic I discuss in Chapter IV. Literature in the economics field continues to search for better explanations of common human behavior that disagrees with theory. Economists have shown that individuals will reduce their wealth in order to achieve fair and/or just outcome in monetary interactions with a stranger.

Additional studies show that individuals are willing to sanction others for not behaving in a fair or just manner. Recent studies conducted by Fehr, Falk and Fischbacher (2008) show that individuals respond to the intentions of a decision, meaning that negative intentions are sanctioned and positive intentions are more likely to be rewarded in comparison to scenarios where intentions are removed from the decision-making process. Here we expand on this concept, by specifically study the impact of attribution of responsibility and two different contexts to see how it impacts the decision-making process. We use "responsibility game" which allows the participants to take responsibility for their decision within the game or place the responsibility on a random device, which makes the decision for them in the experiment. I have explored the affects of direct and indirect responsibility as well as selfish and kind intentions. From this we find that individuals value responsibility of a decision and therefore respond in a different manner than when no responsibility is relevant.

Throughout this dissertation I have employed experimental methodology to collect data on the decision-making process as described in the cases above. From this I have been able to expand on the understanding of the dual-principal agent study and provide simulations of payment mechanisms that I was unable to explore in the laboratory. In this way, I was able to provide insight into the potential outcome of blended payment mechanisms if employed in the dual principal-relationship as studied in the laboratory. Furthermore, I was also able to provide additional insight into the variables which influence human-decision making with respect to responsibility and intentionality.

This dissertation looks at the effect of other-regarding behavior in economic games and who show that it is necessary to include such behavior in economic theory. Through the use of experimental and simulation methods, I have been able to further explore the impact of other-regarding behavior in two different economic scenarios. In Chapters II and III, I study the dual-principal agent relationship. My research shows that other-regarding behavior has a significant impact on the outcome of a contract and therefore classic labor theory does not accurately reflect human behavior in these instances. In the second scenario, I show that by altering the framing of responsibility in an economic game will impact the behavior of individuals participating. Each of these papers show that other-regarding behavior significantly impacts the outcome of these scenarios; therefore, confirming the necessity of addressing this type of behavior in future economic analyses.

CHAPTER II: PAYMENT MECHANISMS IN A DUAL-PRINCIPAL AGENT SETTING: AN EXPERIMENTAL STUDY

ELLEN P GREEN
VIRGINIA TECH
DEPARTMENT OF ECONOMICS

ACKNOWLEDGMENTS: The author would like to thank the Institute for Advanced Study at Virginia Tech for financial support.

1. INTRODUCTION:

Experimental economics is a tool that allows economists to test theories and concepts as well as make inexpensive alterations in procedures, in order to improve the efficiency of procedures prior to implementation in the field. Recently, experimental economists have started using laboratory studies to better understand labor contracts and their relationship to other-regarding behavior, ultimately employing their results in the field. The likely reason for the increase in the popularity of using experimental methodology to analyze labor contracts is that it allows economists to control variables that are impossible to control in the field and therefore to reach more definite conclusions about the consequences of alternative contract designs. With this information firms can vastly improve the efficiency of their infrastructure.

Until now investigations have focused on implications of contract design in the classic principal-agent problem, in which there is only one principal and agents are purely self-interested. This paper, by contrast, uses experimental methodology to examine the effects of contract design on the dual-principal agent problem and the impact of other-regarding behavior on multiple contract designs models.

Analysis of the principal-agent relationship typically focuses on the problem of a single principal and a single agent. This study looks at an underdeveloped topic in economics, which is defined here as the dual-principal agent problem. In the dual-principal agent problem, an agent has responsibilities to two different principals. For instance:

- A lawyer is an agent of clients and also an agent of the lawyer's law firm.
- A physician is an agent of patients and also an agent of the patient's health care provider.
- A professor an agent of the professor's students and also an agent of a university.

For clarity the following language is adopted to describe the actors: the *Employer* is the principal who pays the agent and is also the contract designer, the *Client* is the principal who receives the services of the agent, and the *Agent* is the actor who performs the services. The main focus of this study is payment mechanisms that can motivate agents in cases in which there is a client-specific interior solution for the optimal number of services rendered to the client.

Theoretical papers typically focus on contract design and its implications for the outcome of the interaction between two selfish parties in a classic principal-agent relationship. However, experiments have shown that an individual's decision-making process is directly influenced by other-regarding behavior (Fehr et al. 2007). These experiments also show that other-regarding behavior produces a different optimal payment mechanism for the principal-agent relationship than predicted by theory.

The present analysis uses experimental methodology to extend these findings, investigating the effects of other-regarding behavior on the outcome of multiple payment mechanisms in the context of a dual-principal agent relationship. The payment mechanisms studied in this analysis are Piece Rate, Flat Rate, Salary, Socialization, Flat Rate with Bonus, and Piece Rate with Bonus. These payment mechanisms are the most commonly used ones in dual-principal agent relationships in practice. Thus experiments with them permit us to explore the effect of other-regarding behavior in the context of the dual-principal agent relationship as it occurs in the field, and subsequently to test the capacity of economic theories to optimize contract design. Economic theory suggests that each of these payment mechanisms creates a significantly different incentive structure for agents in dual-principal relationships (Robinson 2001). The Piece Rate payment mechanism, which pays per service rendered, creates an incentive for agents to overprovide services. The Flat Rate payment mechanism, which pays a flat rate per client, encourages agents to underprovide services to each client, but to treat more clients. The Salary payment mechanism provides no monetary incentives for agents to perform services. Bonuses are used to incentivize agents to provide a specific service dictated by the employer. The bonus incentive is often seen in the health care industry as encouraging physicians to meet quotas for specific procedures as determined by the United States Preventative Services Task Force.[1] Quota Bonuses as studied here incentivize agents to overinvest their time in those tasks that are specifically rewarded by

[1] http://www.uspreventiveservicestaskforce.org/about.htm

their completion, leading to a lack of effort spent on other tasks (Peterson, Woodward et al. 2006). The socialization payment mechanism studies the effect that the knowledge of other agents' behavior has on agent effort.

According to classic agency theory each payment mechanism, due to its unique incentives, should result in a different quantity of services and therefore a different level of quality. For instance, labor theorists predict that the Salary payment mechanism will result in the lowest frequency of services and therefore the lowest quality of care, because there is no monetary incentive for agents to provide services (Gosden 2003). However, the data from this experiment show that payment mechanisms that are weakly linked to actions of agents or are *prospective* (i.e. Salary, Flat Rate, and Flat Rate with bonus) result in higher quality outcomes when compared to payment mechanisms with monetary incentives that are directly linked to agent actions, or are *retrospective* (i.e. Piece Rate and Piece Rate with Bonus). This indicates that classical contract theory has failed to model human behavior appropriately with respect to the effects of incentives and, furthermore, has misrepresented the efficiency of various payment mechanisms. Additionally, the results indicate that many economists studying field data in the dual-principal agent relationship may have misinterpreted the quality of services provided in their results.

This study isolates the effects of payment mechanisms on agent behavior in the dual-principal agent relationship by using a novel experimental game, "The Dual-Principal Agent Game." This game captures the important aspects of the dual-principal agent relationship and removes the impact of exogenous variables that often influence outcomes in the field, so that we can make more definite conclusions about the variables that influence agent behavior. The remainder of the paper is organized as follows: Section 2 provides a literature review of contract design and its implications for economic theory and experiments, and in the field; Section 3 describes the experimental design and procedures. Then the results of the study are reported in Section 4, followed by the conclusions in Section 5. Section 6 provides a summary of planned future work and relevant applications of the findings from the experiment. The Appendices provide the materials from the experimental process.

2. LITERATURE REVIEW

This section provides a discussion of the relevant literature on contract design from theory, experiments and field analysis.

A. AGENCY THEORY

Agency theory emphasizes the need for contract designers to directly link incentives to an agent's tasks. The ideal payment structure rewards agents for increasing their principal's utility. However due to imperfect information and costly monitoring, this is nearly impossible to contract in most principal-agent relationships. And while designing a good contract can include creating incentives to increase production, by increasing wages or bonuses, as well as creating incentives to avoid decreasing production, by imposing punishments or cutting wages, many principal-agent relationships require careful balancing of these incentives. In many principal-agent relationships there is an interior optimum for the number of services to be provided to the client, and deviating from this quantity is detrimental to the client. This problem is commonly seen in the health care industry, where overprovision or under-provision of services to a patient (the client) by a physician (the agent) is likely to be both damaging to the patient's health and costly.

In his review of physician payment mechanisms, Robinson (2001) describes the outcomes that the incentives behind commonly used payment mechanism are likely to produce. In particular, he addresses three common mechanisms, which he describes as least aligned to the optimal outcome for the patient. These are the Flat Rate, the Salary, and the Piece Rate payment mechanism.[2] The Flat Rate payment mechanism, by paying a rate per patient, encourages physicians to treat as many patients as possible. However, it also encourages "cream skimming" of patients because it does not adjust appropriately for the additional cost of the time and effort required for the sickest patients (Matsaganis and Glennerster 1994). Thus, rather than bear the cost of sickly patients, physicians select the healthier and more profitable patients. The Salary payment mechanism creates no strong links between physician effort and payment; in theory it should therefore perform poorest in terms of quality of services. Physicians under the Salary payment mechanism are often encouraged through social incentives, promotions, risk of termination and other incentives to provide better patient care. And finally there is the Piece Rate mechanism. While it encourages physicians to do all that is necessary to improve patient well being, it also encourages overtreatment to improve the physician's payoff. Many economists agree that the Flat Rate payment mechanism leads to Physician Induced Demand (PID), defined by Evans (1974) as a shift in a patient's demand curve that is induced by a physician and benefits the physician. Thus physicians paid under the

[2] In the Health Care Literature the Piece Rate payment mechanism is commonly referred to the Fee-For-Service and the Flat Rate is referred to as Capitation. To preserve the consistency in the text I will refer to them throughout as the Flat Rate mechanism and Piece Rate payment mechanism.

Piece Rate payment mechanism have a financial incentive to use PID to increase the demand for services until the physician is satisfied with the outcome, irrespective of the cost or the patient's well being.

In addition to the effect of monetary incentives on agent performance, it is important to note the impact of other-regarding behavior on agent performance. Hideshi (2004), Dur and Glazer (2008), Kragl and Schmid (2009) and Englmaier and Wambach (2003) develop models in which they address the impact of agents' inequity aversion on optimal contract design in the classic principal agent relationship. These models agree that ignoring other-regarding behavior in contract design has a negative impact the outcome of a contract. Taking account of this human propensity permits closer approximations to optimal contract design. The theories explored in these models are based on the findings with respect to other-regarding behavior that have come from experimental economics. I explore this topic in the next section of the literature review.

B. Experimental Analysis

While theorists have discussed both the dual-principal agent problem, and the classic principal-agent problem with inequity aversion, there have been no studies in the theoretical literature addressing the impact on contract outcomes of other-regarding behavior in the context of the dual-principal agent problem. There is a vast literature in experimental economics that shows a significant impact of other-regarding behavior on the decision-making process. For instance, Stahl (1972) showed that individuals value fairness in addition to their own monetary gain. Furthermore, a number of studies, including Guth et al. (1982) and Roth et al. (1991) used experiments involving the ultimatum game to test the existence of other-regarding behavior in decision-making tasks. Each of these studies found that the players were willing to forego a nontrivial amount of money to achieve fairness or equity in the allocation of resources. Following these results, Forsythe et al. (1994) compared the results of both a dictator game and an ultimatum game and showed that people's inclinations to distribute wealth are better explained by altruism than by inequity aversion. Each of these studies produced data that was inconsistent with the Nash Equilibrium predicted by standard game theory for purely self-interested agents. Although the precise motivation (i.e. altruism, inequity aversion, concern for fairness, etc.) behind this behavior is unclear, ignoring the effect of other-regarding behavior on the decision-making process results in a failure to predict human behavior accurately. These findings have encouraged several new areas of research in experimental economics, each trying to further explore the impact of such behavior on economic scenarios.

More recent experiments show that other-regarding behavior has a significant impact on how successfully contracts are implemented in a principal-agent relationship. In Fehr et al. (2007) principals who acknowledged the altruistic nature of their agent when constructing their contracts achieved higher payoffs than principals who modeled their contracts on the assumption that their agents were self-interested. Furthermore, Charness (2010) found that an agent's productivity increased when information regarding the agent's relative productivity in a group was known. In another study addressing agent position and agent effort, Hoffman et al. (1994), found that individuals behave in an increasingly selfish manner when they earn rather than randomly receive the position of the dictator in the “Dictator game.” Their results indicate that the context of the contract alters the outcome and therefore needs to be taken into account in contract design. Fehr et al. (2002) show that the framing of incentives has a significant effect on agent productivity. In his experiments, agents were less productive when incentives were framed as penalties than when they were framed as bonuses. This disparity in agent productivity is inconsistent with consequtialist theory; if the outcome is the same, the subjects should be indifferent. Subjects here showed that positive reinforcement was more readily acceptable and efficient than negative reinforcement.

In light of laboratory experiments on contract design, a field experiment was conducted in British Columbia, Canada. Shearer (2004) explored the productivity effects of two payment mechanisms, Piece Rate and Flat Rate on workers in a tree-planting firm. Here the author found that the Piece Rate mechanism resulted in 38.11% increased production when compared to a Flat Rate payment mechanism. However, this increased rate of production should not necessarily be interpreted as increased quality production. There were issues regarding the quality of the product under the Piece Rate payment mechanism. In particular, the relationship between the agent and the principal had an interior solution, because some locations where seedlings could be planted were desired and others were not. Therefore, an increase in seedlings planted did not always imply an increase in the quality of production, so that one cannot conclude that the Piece Rate mechanism outperformed the Flat Rate mechanism in quality, which is the question of interest to the contract designer.

C. Empirical Studies in the Field

Empirical studies in the field often provide inconsistent conclusions about the relative efficiency of alternative payment mechanism in the dual-principal agent relationship. This is due to the difficulties that often arise in the form of exogenous variables. Demographics, timing, and procedure all have a significant impact on agent behavior and are can be impossible to control in the field without interrupting business practices. However, there is still a significant amount of research in the dual-principal agent relationship in the Health Care Industry, and the conclusions of this research are noteworthy.

In papers by Davidson, Manheim et al. (1992) and Holhen, Manheim et al. (1990) the authors studied the change in physician behavior when paid under Piece Rate versus Flat Rate payment mechanisms. Both studies found no indication that, as predicted by economic theory, physicians paid under a Flat Rate payment mechanism provided fewer services than those under the Piece Rate mechanism. Furthermore, a literature review by Gosden et al. (2003), reported that there is little evidence of differences in the quantity of treatments among all payment mechanisms. However, they noted that in the qualifying studies, the instances in which there were no significant differences in the quantities of treatment, were ones in which the authors had failed to control for possible demographic differences among the treated populations. This failure can potentially change the conclusions of the studies, explaining the lack of differences (Kakwani, Wagstaff et al. 1997). Gosden concluded that although there is evidence that the quantity of care is greater when the Piece Rate payment mechanism is used than when either the Flat Rate or the Salary payment mechanisms are used, the results were not robust enough to justify a significant change in payment policies before more thorough and controlled research is undertaken.

On the other hand, some researchers have found a significant impact on physician behavior caused by payment mechanism, as predicted by economic theory (Hellinger 1996; Hillman, Welch, and Pauly 1992; Gold Nelson, Lake, et al. 1995). However, these authors stipulate that their findings are subject to bias, and that in order truly isolate the impact of the payment mechanisms one must control for variables that are nearly impossible to control in the field without hindering the practices' production.

4. DATA AND METHODOLOGY:

Using experimental methodology this investigation explores [1] the behavior of agents in the newly defined dual-principal agent relationship and [2] how the success of a payment mechanism's incentive structure is influenced by other-regarding behavior. Because there have not yet been any experimental studies that have addressed this specific interaction in the dual-principal agent relationship, the "Dual-Principal Agent game" was developed.

The "Dual-Principal Agent game" differs significantly from the experimental design employed by Fehr et al. in their 2002 and 2007 studies of the standard principal-agent relationship. In the Fehr et al. studies, principals earned a profit equal to the 'work product' created by the agent minus the agent's salary, and the agents earned a profit equal to the salary less the cost of the effort. In their studies, the value of work produced and the cost of effort were numbers read from a table that the subjects selected, requiring no real effort to be exerted. What is different in the "Dual-Principal Agent game" is that agent effort is generated by the actions of agents, meaning that the subjects perform a real-effort task, and their actions in this task determine their final payments. Furthermore, this game does not attempt to quantify the cost of effort to the agents, which is consistent with real work situations, where the employers cannot quantify the costs to their employees (agents) of performing services for clients. However, we do assume that when an agent performs a service, the benefit to the agent outweighs the cost.[3]

In other respects, the framework of the "Dual-Principal Agent game" is similar to that of the traditional, principal-agent game. As in Fehr's studies, the employer or (secondary) principal and the agent in the "Dual-Principal Agent game" interact in a one-time application of their contract. The fact that there is only a single interaction eliminates the influences of agent reputation, promotions, and/or the possibility of contract termination on the outcome of the contract. Additionally, the principals have no authority over which agents they are pared with, eliminating the possibility for the principals to employ any screening mechanisms in their efforts to improve outcomes. Furthermore, The agents are not given the opportunity to select their secondary principals (employers) based on the payment mechanism; this eliminates the possibility of adverse selection. Thus this game allows us to focus exclusively on the effects of the different payment mechanisms on agent behavior.

[3] Time was limited in the agent's real effort task and therefore, we cannot conclude that the converse is true.

A. EXPERIMENTAL DESIGN:

Each participant in the "Dual-Principal Agent game" represents a role in the dual-principal agent relationship. The experimenters function as the employers (secondary principals) by assigning the payment mechanisms and providing payment to the agents. Each subject acts either as the agent or the client in the relationship.

The "Dual-Principal Agent game" takes place in two Phases. In Phase I, each of 10 clients is endowed with $25.00 and asked to complete a proofreading task for 10 essays. Each of the 10 essays has 10 spelling errors that the clients are asked to correct. For each error that the client fails to correct a penalty of $0.25 USD is taken from the client's endowment.[4] The number of spelling errors that are not corrected represents the optimal number of services that the client can receive from an agent.

In Phase II of the game, each subject representing an agent provides proofreading assistance to the ten clients from Phase I of the game. The agent's task is to correct the errors that were missed by the clients in Phase I. To assist the agents in the proofreading task, each error is highlighted in yellow, just as a practicing physician would be more capable of determining an ailment, the agents in the experiment have an advantage in finding errors. However, for each error that was missed by a client in Phase I of the game, the agent saw three yellow highlighted words within the essay. Only one highlighted word was the error missed by the client; the other two highlighted words were included to permit overprovision of services to clients. Each change by an agent to a highlighted word was counted as a service provided.[5] In each session the agents are assigned to a different payment mechanism, as determined by their employer (the experimenter) prior

[4] Note that if the subjects were unable to identify any of the errors, they received no payment.

[5] A change to the text did not have to be correct for an agent to be paid for their services provided.

to the experimental session. At the end of their task, the agents are paid for the changes made, according to the conditions of the agent's particular payment mechanism.

In the "Dual-Principal Agent game" the payment mechanism is the treatment variable. The treatments are Piece Rate, Piece Rate with Bonus, Flat Rate, Flat Rate with Bonus, Socialization, and Salary. Piece Rate agents are paid a set rate for each error they purported to correct, Flat rate agents are paid a set rate for each client they assisted, and Salary agents are paid a set rate for participating in the experiment (refer to Table I for rates). The other three treatments employ one of the above payment mechanisms as a base and add additional incentives to create final payment mechanism. The Socialization payment mechanism pays the agents by Flat Rate payment mechanism and adds the condition that throughout the proofreading task, a table with the behavior of the other agents is available on the screen. Thus we observe the effect of fellow agent behaviors in the experiment on an agent's actions. In the other two payment mechanisms, Flat Rate with Bonus and Piece Rate with Bonus, the agents are given a bonus if they are able to meet a predetermined service quota. In the game, the quota services are identified in the text by highlighting specific errors in the text in a unique color, green instead of yellow. (In the health care industry physicians are often given a bonus if they provide a specified number of services such as mammograms or colonoscopies (Armour, Friedman, et al 2004)). In the experiment, the agents operating under the Piece Rate with Bonus and Flat Rate with Bonus mechanisms earn a bonus of $1.00 if they edit at least nine of the green errors. The agents do not earn the bonus for correcting any quantity of yellow errors. The bonus is added to the earnings as specified by the agent's payment mechanism (i.e. Piece Rate or Flat Rate).

The clients in the game are truly aided by the agents, as they would be in a traditional setting of a dual-principal agent relationship. For each client error that an agent correctly identifies, the client is given $0.15 of the $0.25 lost back after Phases II is completed. To reflect the cost to a client of being overprovided with services, each instance of an agent correcting a mistake in an incorrect manner decreases the client's earnings by $0.05.[6,7]

Table I: Payment Mechanism

Salary	Receives $20.00 at the end of the experiment
Flat Rate	Receives $2.50 for each client assisted
Flat Rate with Bonus	Receives $2.50 for each client assisted, plus a bonus $1.00, if conducted at least 9 of the quota errors
Piece Rate	Receives $0.20 for each highlighted section of the text changed
Piece Rate with Bonus	Receives $0.20 for each highlighted section of the text changed, plus a bonus $1.00, if conductedat least 9 of the quota errors
Socialization (Flat Rate Base)	Receives $2.50 for each client assisted and every peiod a table reflecting the other subject's behavior was displayed in the computer interface

B. PROCEDURE:

The experiments were conducted in the Economics Research Laboratory at Virginia Tech. Students were recruited to participate from Principles of Economics courses offered at Virginia Tech. A total of 136 subjects participated in the experiment. Several sessions were held for both Phase I and Phase II of the "Dual-Principal Agent game".

Phase I of the experiment was completed in the first 3 experimental sessions. In this phase, recruited subjects played the role of the client in the dual-principal agent relationship. In each session, as the participants entered the laboratory lobby they were

6 Agents were not penalized for making incorrect changes to errors.

7 If the case that the agent did more harm by editing the clients paper than help, the client was not required to pay additional monies.

asked to complete a consent form and wait for further instruction. Once all the subjects arrived, the subjects were escorted into the laboratory's classroom and asked to edit the 10 essays within a 50-minute timeframe. The subjects were spaced so that they could not view the other participants' actions in the task. Subjects were free to leave prior to the allotted time if they were satisfied with their effort. Once the task was completed, subjects were paid individually according to their success in completing the task. At this time, the subjects were also given further information about how some of their endowment would be earned back through assistance of participants in Phase II of the experiment. If the subjects wished to collect this reimbursement, contact information was collected at this time. Between Phase I and Phase II of the experiment, the proofread essays from the subjects in Phase I were entered into a computer database to be employed in Phase II of the experiment.

In Phase II of the game, different subjects were recruited to represent the agents in the dual-principal agent game. This phase was conducted over 13 sessions. In each of these sessions the subjects entered the laboratory lobby and were asked to fill out a consent form and wait for further instruction. Once all the subjects were present they were escorted into the computer laboratory and asked to sit at a terminal. Once seated, the subjects were given instructions for their task. The details of the payment mechanism were provided to the subjects both verbally and in text in each session. At this time subjects were given the opportunity to ask questions about their task. Once the instructions were complete and all questions had been answered, the agents were asked to complete a quiz on the details of their payment mechanism to ensure that they fully understood how their actions would affect both their payoff and the client's payoff (See

Appendix B for quiz). The experiment continued only when all subjects were able to successfully answer the quiz questions.

Next, the subjects entered a computer interface to perform the proofreading task. The subjects were presented with a screen that allowed them to select any of the 10 essays provided to them, the subjects were able to go back and forth from this screen and each essay. A different subject (client) from Phase I of the game had edited each essay presented to the subjects in Phase II of the game. In this way, the Phase II subjects could potentially assist 10 distinct subjects from Phase I. The interface allowed Phase II subjects to determine how many and which of the Phase I subjects in the essay pool they would provide services for; it was possible for the agent to help none or all of the Phase I subjects if they wished.

To prevent proofreading out of boredom, the subjects were allowed to leave the proofreading task early by exiting the computer interface. After the subjects were finished with their proofreading they completed a brief survey that collected data about their trustworthiness, altruism, reciprocity and socioeconomic status.[8] The agents were then called out to the lobby and paid individually. All of the data from Phase II of the game was collected in order to determine the total reimbursement the participants in Phase I of the game qualified for. At this point, the clients from Phase I were contacted via email to receive their additional earnings.

[8] The altruism, trustworthiness, reciprocity and socioeconomic data collected was used in regression models to test the influence on agent behavior; however none of these variables proved to have a significant impact on agent behavior.

Table II: Summary of Agent's Actions by Client and Treatment

	Client 1	Client 2	Client 3	Client 4	Client 5	Client 6	Client 7	Client 8	Client 9	Client 10	All Clients
All Data											
Harmful	1.13	2.57	3.33	0.45	1.42	0.84	1.72	1.52	2.15	1.86	17.00
Beneficial	1.85	6.28	6.18	0.48	3.44	1.52	2.71	2.79	2.79	2.37	30.43
Total	2.98	8.85	9.52	0.94	4.86	2.37	4.44	4.32	4.94	4.23	47.42
Hurt	37.87%	29.06%	35.03%	48.31%	29.25%	35.57%	38.82%	35.29%	43.50%	43.90%	35.85%
Net	1.20	4.35	4.00	-0.05	2.85	1.15	2.30	2.40	3.10	2.75	24.05
Piece Rate											
Harmful	1.53	3.16	4.68	0.63	1.89	0.89	1.68	1.58	1.84	2.42	20.32
Beneficial	2.00	7.37	6.84	0.26	3.84	1.63	2.79	3.32	3.05	1.89	33.00
Total	3.53	10.53	11.53	0.89	5.74	2.53	4.47	4.89	4.89	4.32	53.32
Hurt	43.28%	30.00%	40.64%	70.59%	33.03%	35.42%	37.65%	32.26%	37.63%	56.10%	38.11%
Net	0.47	4.21	2.16	-0.37	1.95	0.74	1.10	1.74	1.21	-0.53	12.68
Piece Rate with Bonus											
Harmful	4.43	10.48	13.33	0.62	5.81	3.33	7.76	6.86	9.29	7.14	69.05
Beneficial	1.90	6.10	6.14	0.33	2.81	1.10	1.86	2.14	2.29	2.05	26.71
Total	6.33	16.57	19.48	0.95	8.62	4.43	9.62	9.00	11.57	9.19	95.76
Hurt	69.93%	63.22%	68.46%	65.00%	67.40%	75.27%	80.69%	76.19%	80.25%	77.73%	72.10%
Net	-2.52	-4.38	-7.19	-0.29	-3.00	-2.24	-5.90	-4.71	-7.00	-5.10	-42.33
Flat Rate											
Harmful	0.26	0.27	0.61	0.43	0.22	0.13	0.17	0.17	0.70	0.43	3.39
Beneficial	1.83	6.13	5.96	0.52	3.52	1.87	2.91	3.04	3.43	3.52	32.74
Total	2.09	6.40	6.56	0.96	3.74	2.00	3.09	3.22	4.13	3.95	36.13
Hurt	12.50%	4.20%	9.26%	45.45%	5.82%	6.52%	5.63%	5.41%	16.86%	10.99%	9.39%
Net	1.57	5.86	5.35	0.09	3.30	1.74	2.74	2.87	2.73	3.09	29.35
Flat Rate with Bonus											
Harmful	0.29	0.57	0.81	0.24	0.29	0.38	0.29	0.24	0.57	0.57	4.24
Beneficial	1.95	5.90	5.95	0.71	3.29	1.48	2.52	2.48	2.71	2.24	29.24
Total	2.24	6.48	6.76	0.95	3.57	1.86	2.81	2.71	3.28	2.81	33.47
Hurt	12.77%	8.82%	11.97%	25.00%	8.00%	20.51%	10.17%	8.77%	17.41%	20.34%	12.66%
Net	1.67	5.33	5.14	0.48	3.00	1.10	2.24	2.24	2.14	1.67	25.00
Salary											
Harmful	0.23	0.59	0.59	0.36	0.23	0.23	0.14	0.23	0.27	0.23	3.09
Beneficial	2.05	7.18	7.77	0.64	4.05	1.73	3.41	3.18	1.86	1.18	33.05
Total	2.98	8.85	9.52	0.94	4.86	2.37	4.44	4.32	4.94	4.23	36.13
Hurt	7.63%	6.68%	6.21%	38.83%	4.68%	9.61%	3.07%	5.26%	5.52%	5.37%	8.55%
Net	1.82	6.59	7.18	0.27	3.82	1.50	3.27	2.95	1.59	0.95	29.95
Socialization											
Harmful	0.15	0.65	0.35	0.45	0.25	0.15	0.45	0.20	0.35	0.55	3.55
Beneficial	1.35	5	4.35	0.4	3.1	1.3	2.75	2.6	3.45	3.3	27.6
Total	1.50	5.65	4.70	0.85	3.35	1.45	3.20	2.80	3.80	3.85	31.15
Hurt	10.00%	11.50%	7.45%	52.94%	7.46%	10.34%	14.06%	7.14%	9.21%	14.29%	11.40%
Net	1.20	4.35	4.00	-0.05	2.85	1.15	2.30	2.40	3.10	2.75	24.05

5. RESULTS

A. AGENT ACTIONS

A summary of the agents' actions in the "Dual-Principal Agent game" is provided in Table II. Here the agent's actions are assigned to one of two categories: beneficial or harmful. The number of harmful actions is defined by the number of incorrect changes the agent makes to the essays, which resulted in a decrease in the client's payoff of $0.05 per action. The number of beneficial actions is defined as the number of correct amendments to the text the agent makes and therefore results in an increase of the client's payoff by $0.15 per action. Total services is the sum of beneficial actions and harmful actions. Net actions are calculated by subtracting the number of harmful services from the number of helpful services. These variables provide us with a summary of overall behavior induced by each payment mechanism and are reported in Table II as averages over all agents in each payment mechanism.

Table III: Regression with Total Services as the dependent variable and no Constant

Flat Rate	36.13***
	-7.09
Flat Rate with Bonus	33.48***
	-6.28
Piece Rate	53.32***
	-9.51
Piece Rate with Bonus	95.76***
	-17.96
Socialization	31.15***
	-5.7
Salary	36.14***
	-6.94
N	126.00

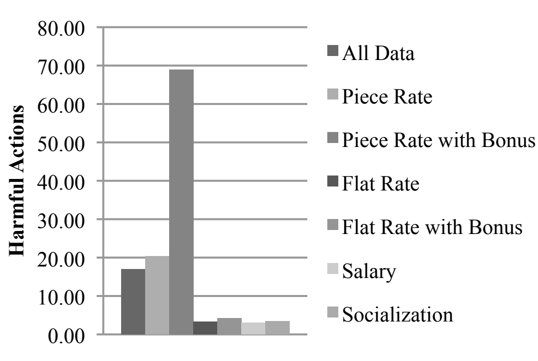
t statistics in parentheses
* p<0.05, **p<0.01, *** p<0.001

Here, the relative total number of services provided by the agents agrees with the predictions of economic theory. Agents paid by the prospective payment mechanisms, the Piece Rate with Bonus and the Piece Rate, provided the greatest number of services. Both of these payment mechanisms encourage agents to increase the number of services they provide by directly rewarding the agents for each service rendered and we see that these incentives successfully influence the agent's behavior. Conversely, the four retrospective payment mechanisms, Salary, Flat Rate, Flat Rate with Bonus and Socialization, provided the fewest number of services in the experiment. Each of these retrospective mechanisms has weak ties between the services provided and the agent's income and therefore as predicted, the agents provide fewer services under these payment mechanisms. An F-test on an least squares regression on total services dependent on payment type (as seen in Table III), reveals a highly significant difference between the mechanisms that provided the greatest number of services and the least number of services (all of the relevant p values were less than 0.03).

In the "Dual-Principal Agent game" the value of services provided to the client is clearly defined by services that increase the client's payoff or the number of beneficial services as presented in Table II. In Table II, we see that agents paid by the Salary payment mechanism provided the greatest number of beneficial services on average to their clients and Piece Rate with Bonus provided the fewest number of beneficial services on average. This behavior can also be seen in Figure 1a.

Figure 1a: Beneficial Actions

Figure 1b: Average Harmful Actions



Classic theory predicts that agents paid by prospective payment mechanisms will provide the fewest number of services as well as spend the least amount of time possible providing services because there is no direct link between number of services and agent income in the Salary payment mechanism. For instance, agents paid by the Flat Rate payment mechanism should theoretically provide the minimum number of services required to obtain a client and leave the experiment once they have done so. However, in this study agents paid by the Flat Rate, Flat Rate with Bonus and Socialization payment mechanisms provided more than the minimum number of services and therefore did not act as purely selfish agents. By running an ordinary least squares regression on the number of harmful services provided we see that the Flat Rate, Flat Rate with Bonus, Salary, and Socialization payment structures all provided fewer than average harmful actions over all clients (as seen in Table IV). On the other hand, the two most harmful payment mechanisms are the Piece Rate and Piece Rate with Bonus for all clients.

Running an F-test on the ordinary least squares regression we find that the Piece Rate with Bonus resulting in the most frequent provision of harmful services to the client.[9] The second most harmful payment mechanism is the Piece Rate, which provided a significant number of more harmful services to the client than the Salary (p=0.0247), Socialization (p=0.0324), Flat Rate (p=0.0258), and Flat Rate with Bonus (p=0.0378) payment mechanisms. Subjects in the "Dual-Principal Agent game" behaved in ways that suggested an alternative payment mechanism, such as Salary or Flat Rate mechanism would reduce wasted spending. This is because the monetary incentives provided by the Piece Rate mechanisms outweighed the human tendency to regard others in their

[9] P-values of F test when compared with Piece Rate with Bonus: Flat Rate (0.00), Flat Rate with Bonus (0.00), Piece Rate (0.00), Salary (0.00), Socialization (0.00).

decision-making process. Therefore, the clients were worse off in the Piece Rate and Piece Rate with Bonus mechanisms than the other mechanisms because of the direct link between services and agent income as observed by the increase in harmful actions.

Table IV: Ordinary Least Squares regression with Harmful Services as dependent variable and no constant

Flat Rate	3.391
	-0.67
Flat Rate with Bonus	4.238
	-0.8
Piece Rate	20.32***
	-3.66
Piece Rate with Bonus	69.05***
	-13.08
Salary	3.55
	-0.66
Socialization	3.091
	-0.6
N	126

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

B. IMPACT ON CLIENT

The impact the agent's actions have on the client is of the greatest concern from both the point of view of the client and policy makers. Ideally, each of the agent's actions positively impact the client's outcome from both the client's perspective and an efficiency perspective. Here, each agent's behavior is tracked based on his or her performance in the proofreading task as described above by harmful and beneficial actions. Impact on the client is calculated by adding +0.15 for each correctly identified error (beneficial actions) and -0.05 for each incorrectly identified error (harmful actions).[10] The client impact variable reflects the direct income effect of the agent's services on the client in the game. This is reported in Table V by the average impact the agent had on the clients' payoff by treatment type. Here the Salary and Flat Rate resulted in the highest impact rating, with an average of $4.80 and $4.74 respectively.

[10] Subjects did not need to correctly identify errors to be paid for their changes, so although these actions were harmful to the clients, under some mechanisms harmful actions were beneficial to the agent.

With these observations, we can comment on the quality of services provided to the client relative to the frequency of services. While in theoretical and field studies authors discuss the decrease in the quantity of services provided when the Flat Rate or Salary mechanisms are employed; they do not fully address the relationship between frequency and beneficial services or harmful services because it is difficult to estimate this value. As shown in the discussion above, there is no significant decrease in beneficial services by agents paid by the Salary or Flat Rate mechanisms due to the observed other-regarding behavior. And although agents were not directly incentivized to provide as many services as they were under the Piece Rate mechanisms, they did not fail to do provide beneficial in these prospective payment mechanisms. These phenomena resulted in an increased benefit to the clients.

Table V: Summary of Agent's Impact on Client by Essay and Treatment

	All Data	Piece Rate	Piece Rate w. B	Flat Rate	Flat Rate w. B	Salary	Socialization
All Data							
Average Negative Impact	$0.85	$1.02	$3.45	$0.17	$0.21	$0.15	$0.18
Average Positive Impact	$4.56	$4.95	$4.01	$4.91	$4.39	$4.96	$4.14
Net impact	$3.71	$3.93	$0.55	$4.74	$4.17	$4.80	$3.96
N	126	19	21	23	21	22	20

C. CLIENT SELECTION

Logit regressions are used to determine if the probability of a client receiving services from an agent was dependent on the payment mechanism. The results are provided in Table VI. Here, agents paid in the Flat Rate, Flat Rate with Bonus and Socialization mechanisms had a higher probability of providing proofreading aid for clients 4, 8, 9, and 10 than the clients paid via Piece Rate and Piece Rate with Bonus. This first result, for clients 4, 8, 9, and 10 is predicted by agency theory. In Robinson's (2001) discussion of the effects of payment mechanism on agent behavior, he comments on the idea that those agent's paid with the Flat Rate based mechanism will help more clients than those paid via mechanism with no direct monetary link to the number of clients they assist. Clients 5 and 6 are more likely to be provided services by the Piece Rate payment mechanism and Piece Rate with Bonus Payment mechanism. It is counterintuitive that the Flat Rate payment mechanism would be less likely to provide services for any of the clients, because their income increased with each client they served. Overall, the Flat Rate payment mechanism payment mechanism resulted in the most clients served, but not consistently on a per client basis.

Table VI: Likelihood of Client Selection by Payment Structure from Logit Regression

	Flat Rate	Flat Rate with Bonus	Piece Rate	Piece Rate with Bonus	Sociailization	Salary
Client 1	100%	100%	100%	100%	100%	95%
	-	-	-	-	-	(0.004)
Client 2	100%	100%	100%	100%	100%	100%
	-	-	-	-	-	-
Client 3	100%	100%	100%	100%	100%	95%
	-	-	-	-	-	(0.004)
Client 4	80%	78%	60%	78%	85%	100%
	(1.20)	(1.20)	(0.98)	(1.20)	(0.63)	-
Client 5	100%	50%	100%	95%	100%	100%
	-	(1.45)	-	(1.02)	-	-
Client 6	52%	50%	100%	95%	100%	100%
	(1.45)	(1.45)	-	(1.02)	-	-
Client 7	100%	17%	46%	22%	100%	95%
	-	(1.16)	(1.45)	(1.19)	-	(1.02)
Client 8	62%	40%	46%	40%	75%	86%
	(0.97)	(0.83)	(0.88)	(0.83)	(1.20)	(0.62)
Client 9	100%	76%	73%	76%	95%	50%
	-	(0.67)	(0.67)	(0.67)	(1.11)	(0.43)
Client 10	95%	81%	66%	74%	94%	36%
	(0.86)	(0.66)	(0.64)	(0.63)	(0.87)	(0.44)

D. IMPACT ON EMPLOYER

In application, employers, or the contract designer, are most interested in how payment mechanisms affect their firm both in the quality of product they offer and the cost of employing the payment mechanism. The only outright cost in the "Dual-Principal Agent game" is the cost of hiring the agent in the proofreading task. In other words, in the game the total cost was the cost to the experimenter for hiring the subject; however, there was no direct benefit received by the employer for hiring the agent in the experimental design. In industry the benefit to the employer should be reflected both by the quality of services provided to the client and they payment from the client from these services.

Table VII: Costs to employer by Payment Type

	All Data	Piece Rate	Piece Rate w. B	Flat Rate	Flat Rate w. B	Salary	Socialization
Costs	$21.33	$10.77	$20.18	$24.35	$22.81	$25.00	$23.50
Benefits	$4.56	$4.95	$4.01	$4.91	$4.39	$4.96	$4.14
Benefits/Cost	0.22	0.49	0.13	0.19	0.18285	0.19	0.17
N	126	19	21	23	21	22	20

First, consider the cost to the employer for hiring the agent. Table VII reports the average cost to the employer for hiring an agent by treatment type. The highest cost to the employer was the Salary ($25.00) payment mechanism, followed closely by the Flat Rate payment mechanisms ($24.35).[11] This result is surprising given that the incentives provided in the Piece Rate and Piece Rate with Bonus mechanisms create a direct relationship between the income of the agent and the number of services provided; therefore, allowing the agents to maximize their income (thereby increasing their cost to the employer) by changing as many errors as possible.[12] However, the agents paid by these mechanisms did not conduct enough services to increase their cost to the point of the other mechanisms. This result could potentially be an artifact of the time limit on the game.[13]

Second, employers are concerned with the ratio of the cost of services to the quality of services provided. Although, not directly studied in this experiment, the reputation of a firm can significantly impact future revenues (McGuire 2000). Although employers can potentially keep costs low by employing the Piece Rate payment mechanism, it is important for them to realize the impact that the quality of services has

[11] The Salary and Flat Rate payment mechanism was significantly higher than the Piece Rate and Piece Rate with bonus; the relevant p-values in comparison were all less than .01.

[12] One of the difficulties in creating the wages for each payment mechanism was in trying to create values that did not directly affect the agent behavior, however still provided incentive to reveal the incentives of each payment mechanism. Therefore, I hesitate to solely compare the cost estimates and think it is more efficient to focus on the cost/benefit analysis provided in the next section.

[13] The time constraint on the proofreading task could have potentially influenced this, however there were subjects that did change almost all of the errors available for them to change in the given time frame.

on future business. Many dual-principal agent relationships involve multiple interactions between the principal and agents before all of the services necessary have been provided. Therefore, reputation and quality of services provided during the interactions can seriously impact whether or not a client will be a repeat or a single time user. In the long run, focusing on quality over initial cost is important.

Table VII compares the relationship between quality of product and the cost of employing the payment mechanism. Here, the benefit-to-cost ratio is the net difference of actions (beneficial-harmful services) to the cost of employing the client.[14] Although a significant amount of money was wasted on services provided in the Piece Rate payment mechanism, the number of beneficial services provided outweighs the waste resulted in the highest benefit-to-cost ratio. This benefit-to-cost ratio for the Piece Rate mechanism was significantly higher than all other payment mechanisms. The expected payoff to the agent in the Piece Rate mechanism was significantly lower despite the poor quality of services;[15] however the cost-benefit analysis showed that this mechanism provided the highest benefit-to-cost ratio. Conversely, the Piece Rate with Bonus continued to fare the worst and produced results with significantly lower quality in the game when costs are compared to benefits. Although the differences are not highly significant, the results indicate the following ranking in the cost benefit analysis: Piece Rate, Salary, Flat Rate, Socialization, Flat rate with Bonus, and Piece Rate with Bonus. Again, emphasis should be placed on the differences amongst the costs and benefits within each industry.

[14] It is important to note that the only cost to the employer in the industry setting is not entirely collected in the payment of the agent. Employers must provide many fixed-cost for the agent to work at their firm. Additionally, the Health Care Industry must purchase Malpractice Insurance. These additional costs can increase the cost of the overprovision of services, so the cost of Piece Rate mechanism in this variable is perhaps underestimated.

[15] P Values of Piece Rate vs other payment mechanisms: Flat Rate (0.00), Flat Rate with Bonus (0.00), Piece Rate with Bonus (0.00), Salary (0.00), and Socialization (0.00).

6. Conclusions and Relevant Applications

The results of this study provide a novel vantage point for contract designers to approach the optimal payment mechanism. The behavior observed by the agents in the "Dual Principal Agent game" contradicts the predictions labor theory. Here, it was shown that those payment mechanisms that created a direct link between agent income and number services provided resulted in the lowest frequency of quality services, i.e. Piece Rate and Piece Rate with Bonus. However, those payment mechanisms that were loosely tied to actions or number of services provided resulted in a higher frequency of quality services, i.e. Salary, Flat Rate, and Flat Rate with bonus that can be attributed to the existence of other-regarding behavior. In terms of service frequency, the experiment resulted in findings as is predicted with contract theory. The Piece Rate and Piece Rate with Bonus mechanisms provided a greater number of services than the Flat Rate, Flat Rate with Bonus, Salary, and Socialization mechanisms. Additionally, the cost benefit analysis showed that the highest quality for the cost was provided by the Piece Rate payment mechanism.

One of the direct applications of these results is within the Health Care Industry. Spending on health care continues to outpace the Nation's gross domestic product, which threatens the sustainability of the system. Approximately 52% of health care spending can be attributed to hospital care and physician or clinical services. Due to its significant portion of the spending in Health Care, it is important to understand from a policy maker's perspective the portion of those services that are wasteful and which are beneficial. In this investigation, the harmful services provided in the Piece Rate with Bonus payment mechanism were 72.10% of all services. If we assumed that this percentage reflected the waste observed in health care spending at least 37.5% of total health care spending is spent on wasted or detrimental services rendered. Furthermore from the patient's perspective, the Organization for Economic Co-Operation and Development (2010) estimates that average spending per capita on health care in a year is $7,538. To put this in perspective based on the results of the experiment, a physician being paid under the Piece Rate with Bonus mechanism would provide $4,949.93 worth of harmful services and only $2,588.07 of beneficial services.[16] This number does give an indication of the dollars wasted based on one of the predominantly used payment structures. Based on the results of this study, the Salary payment mechanism would result in a decreased amount of waste in the health care industry, however further investigation is required before such a change is made.

The "Dual-Principal Agent game" shows how payment mechanisms are affected by other-regarding behavior and while these findings provide a new perspective on contract design, there is much research that still needs to be conducted prior to implementation, and there are many other aspects of contract design that need to be addressed in future experiments.

7. Future Work

While these experiments show how agent behavior in the dual-principal agent relationship is affected by different payment mechanisms, there is still much work needed to be able to fully understand why these results exist in a laboratory setting and not in the field. Many dual-principal agent relationships are in need of a solution to their poorly

[16] This number is only based on the average total spending in the United States per capita and all physicians are not paid under the Piece Rate with Bonus payment structure.

incentivized payment mechanisms and conducting further experiments could potentially reveal ways to improve contract design. Future experiments need to address risk, novel payment mechanisms, and multi-period interactions.

Risk is has a significant impact on the human decision-making process and ignoring the influence of risk can greatly weaken the predictability of a model and needs to be addressed in future investigations of the dual-principal agent relationship. Many dual-principal agent relationships in the field result in lawsuits based on poor quality of service. For instance, in the Health Care Industry physicians are often sued for malpractice, which creates an additional cost to providing harmful services to both the Health Care Provider and the physician. The "Dual-Principal Agent game" does not allow the client to react to the services provided by their agents. Adding the ability to fine or punish agents for unsatisfactory behavior could increase their quality of work due to added risk. Therefore, this cost can potentially change the outcomes of the game and should be taken into account in further studies.

Each of the prospective and retrospective payment mechanisms studied here create incentives that successfully address the need for the principal, while at the same time encouraging inefficiencies in services provided. The Flat Rate, encourages cost efficiency through risk-sharing, however the extent to which this is present can potentially result in principals being underprovided for. The Piece Rate encourages agents to provide as many services as necessary rather than the under provision of services, however allows the agent to profit off of the over provision of services. In many applications of the dual-principal agent relationship a blend of these mechanisms have been employed. The success of these blended mechanisms is difficult to determine due to the exogenous variables observed in the field. While these experimental show the investigator how agents behave in these base payment mechanisms, further experimental exploration into blended mechanisms and novel mechanisms. These potential experiments will provide further insight to understanding how the optimal payment mechanism should be designed in this relationship.

Finally, this experiment focused on the effect of other-regarding behavior in a dual-principal agent relationship over a one-time contract. Altering the number of interactions between a client and their agent and employer would potentially alter the outcome of the contract. For instance, how agents interact with the same client could evolve to a more selfish relationship as is the case in previously conducted multiple session experiments (Guth 1993). However, McGuire (2000) models repeated interactions between a physician, their patient, and their employer and notes that the net benefit of the patient affected the market demand for the physician's services therefore a physician should acknowledge the signal they send through their effort to attract potential and repeat clients. In this way, if one allows the clients to select their agents in each interaction, this could result in potential signaling to the client through agent effort. Each of these factors is viewed in field applications of the dual-principal agent relationships and therefore need to be studied in future experiments. In conclusion, while this investigation introduces new aspects to contract design, there are ample opportunities for future researchers to explore the dual-principal agent relationship.

REFERENCES:

Armour, B. S.; C. Friedman; M. M. Pitts; J. Wike; L. Alley and J. Etchason. 2004. "The Influence of Year-End Bonuses on Colorectal Cancer Screening." *American Journal of Managed Care*, 10(9), 617-24.

Bolton, P. and M. Dewatripont. 2005. Contract Theory. Cambridge, Mass.: MIT Press.

Bull, C.; A. Schotter and K. Weigelt. 1987. "Tournaments and Piece Rates - an Experimental-Study." *Journal of Political Economy*, 95(1), 1-33.

Charness, G.; D. Masclet; and M. Villeval. 2010. "Competitive Preferences and Status as an Incentive: Experimental Evidence." *IZA Discussion Paper No. 5034*

Davidson, S. M.; L. M. Manheim; S. M. Werner; M. M. Hohlen; B. K. Yudkowsky and G. V. Fleming. 1992. "Prepayment with Office-Based Physicians in Publicly Funded Programs - Results from the Childrens Medicaid Program." *Pediatrics*, 89(4), 761-67.

Dur, R. and A. Glazer. 2008. "Optimal Contracts When a Worker Envies His Boss." *Journal of Law Economics & Organization*, 24(1), 120-37.

Englmaier, F. and A. Wambach. 2010. "Optimal Incentive Contracts under Inequity Aversion." *Games and Economic Behavior*, 69(2), 312-28.

Evans, R. 1974. "Supplier-Induced demand: some empirical evidence and implications." *The Economics of Health and Medical Care.* 162-173

Fehr, E.; U. Fischbacher and S. Gachter. 2002. "Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms." *Human Nature-an Interdisciplinary Biosocial Perspective*, 13(1), 1-25.

Fehr, E.; A. Klein and K. M. Schmidt. 2007. "Fairness and Contract Design." *Econometrica*, 75(1), 121-54.

Feldman, R. 1988. "Competition among Physicians, Revisited." *Journal of Health Politics Policy and Law*, 13(2), 239-61.

Forsythe, R.; J. L. Horowitz; N. E. Savin and M. Sefton. 1994. "Fairness in Simple Bargaining Experiments." *Games and Economic Behavior*, 6(3), 347-69.

Frydman, C. and D. Jenter. 2010. "Ceo Compensation." *Annual Review of Financial Economics*, Vol 2, 2, 75-102.

Gosden, T.; B. Sibbald; J. Williams; R. Petchey and B. Leese. 2003. "Paying Doctors by Salary: A Controlled Study of General Practitioner Behaviour in England." *Health Policy*, 64(3), 415-23.

Guth, W.; P. Ockenfels and M. Wendel. 1993. "Efficiency by Trust in Fairness - Multiperiod Ultimatum Bargaining Experiments with an Increasing Cake." *International Journal of Game Theory*, 22(1), 51-73.

Guth, W.; R. Schmittberger and B. Schwarze. 1982. "An Experimental-Analysis of Ultimatum Bargaining." *Journal of Economic Behavior & Organization*, 3(4), 367-88.

Gurtler, M. and O. Gurtler. 2011. "Inequality Aversion and Externalities." *SSRN eLibrary*.

Hoffman, E.; K. Mccabe; K. Shachat and V. Smith. 1994. "Preferences, Property-Rights, and Anonymity in Bargaining Games." *Games and Economic Behavior*, 7(3), 346-80.

Holmstrom, B. and J. Tirole. 1991. "Transfer Pricing and Organizational Form." *Journal of Law Economics & Organization*, 7(2), 201-28.

Itoh, H. 2004. "Moral Hazard and Other-Regarding Preferences." *Japanese Economic Review*, 55(1), 18-45.

Jacoby, M. B.; T. A. Sullivan and E. Warren. 2001. "Rethinking the Debates over Health Care Financing: Evidence from the Bankruptcy Courts." *New York University Law Review*, 76(2), 375-418.

Kragl, J. and J. Schmid. 2009. "The Impact of Envy on Relational Employment Contracts." *Journal of Economic Behavior & Organization*, 72(2), 766-79.

Lazear, E. P. and S. Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, 89(5), 841-64.

Ma, C. T. A. and T. G. McGuire. 1997. "Optimal Health Insurance and Provider Payment." *American Economic Review*, 87(4), 685-704.

Mcguire, T. G. 2000. “Physician Agency.” The Handbook of Health Economics. Elsevier, 1:1(1). 462-536.

Petersen, L. A.; LeChauncy D. Woodard; Tracy Urech; Christina Daw and Supicha Sookanan. 2006. "Does Pay-for-Performance Improve the Quality of Health Care?" *Annals of Internal Medicine*, 145(4), 265-72.

Rice, T. H. 1983. "The Impact of Changing Medicare Reimbursement Rates on Physician-Induced Demand." *Medical Care*, 21(8), 803-15.

APPENDIX A:

Introduction

This is an experiment in decision-making. You will have an opportunity to earn a considerable amount of cash through your participation in this experiment. You will be required to complete the experiment individually. You are not allowed to communicate with any other participant at any point during the course of the experiment. If you complete the experiment, you will be paid for the decisions you make in the task. You may withdraw from the experiment at any time. If you withdraw from the experiment before its completion, you will only be paid for the portion of the experiment that you have completed. If you do remain in the experiment, then you should feel free to try to make as much money as you can. All responses and decisions will be anonymous and the only piece of experimental material that will contain your identity will be your receipt of payment at the end of the experiment.

Before we begin, please set your cell phones to silent. We ask that you not make calls or send text messages until the experiment is complete. We also ask that you not talk to other participants in the experiment until after the experiment is complete.

Instructions

Basic Overview:
In this experiment you will be asked to proofread 10 short essays. The number of essays that you proofread and the number of corrections per essay that you complete is up to you. Your income earned in this experiment will be based on your performance in the task.

Proofreading Task:
In a previous session of this experiment we asked participants to proofread short essays for typographical and spelling errors. They were given 10 essays with 10 errors in each essay. These participants were initially given $25.00 for completing the proofreading task, but had $0.25 taken away for each error they were not able to correct.
For instance:

- If in Essay 1, the participant corrected 5 out of 10 errors, $0.25*5=$1.25 was taken away from their initial endowment.
- If in Essay 2 the participant corrected 0 out of 10 errors, $0.25*10=$2.50 was taken away from their initial endowment.
- A total of $3.75 was taken away from their initial endowment in this example.

Note that if the participant was unable to correct any errors, they received $0.00 for the experiment.

Your task in the experiment will be to proofread the essays that the first participants have already edited. The changes that you make to their essays will allow them to earn back a portion of their $25.00 that they lost. For each error that you correctly identify, we will reimburse the participant in the first group $0.15. However, if you make an incorrect change the participant in the first group will lose $0.05.

The mistakes that the first participants identified will no longer be in the essay and therefore you will not need to correct them.

The mistakes that the first group missed will still be in the essay. Your task is to proofread these essays for the errors that the first group missed. You will be given assistance with your editing task. For each error in the essay, we have highlighted 3 words, only one of which is the actual error. The first group was not given this assistance in their proofreading task. For example, if the person who initially proofread the essay missed 3 errors you will see 9 highlighted words, only 3 of which contain actual errors. The total number of actual errors will be reported to you at the top of each essay.

This proofreading task allows each participant from the previous session to obtain help from one person in this part of the experiment for his or her essays.

In total you will be presented with 10 essays to edit, each from a different participant. You will have a total of 10 minutes to complete this task. At the end of the 10-minute period you will no longer be able to proofread essays.

Proofreading Payment:
You will be paid $0.20 for each of the highlighted words that you attempt to correct within the 10-minute time frame. You do not have to correctly change the typo to be paid, but the person whose essay you are editing will only be reimbursed for correct editing and will lose money if incorrect editing.

For instance if you make 9 changes to Essay 1 your payment will be calculated as follows:

- If you correctly identify 3 of the highlighted errors and incorrectly identify 6 of the highlighted errors:
    - You will receive $0.20 for all 9 changes and your total payment for Essay 1 will be $1.80
    - The participant whose essay you are editing will receive $0.15 for each of the 3 correctly identified errors (a total of $0.45) and lose $0.05 for the 6 incorrectly changed words (a total of $0.30), so their total reimbursement payment will be $0.45-$0.30=$0.15.

Note: Your total payment will be calculated by summing your changes over all 10 essays.

Completion of Task:
You will be given 10 minutes to complete the proofreading task, however if you are satisfied with the number of corrections that you have made you can leave the experiment by clicking the “Finish” button on the lower left-hand side of your screen before the 10 minutes are up. At this time the instructor will provide you with a brief survey to complete before you are paid and are free to leave.

APPENDIX B:

Quiz: *You will now complete a quiz to see if you understand how your earnings will be calculated.*

1. If you correctly identify 5 of the highlighted errors and incorrectly identify 1 of the highlighted errors in Essay 1:

For your corrections in Essay 1 you receive _____
The participant whose essay you are editing will receive _____

2. If you correctly identify 1 of the highlighted errors and incorrectly identify 5 of the highlighted errors in Essay 2:

For your corrections in Essay 2 you receive _____
The participant whose essay you are editing will receive _____

3. If you correctly identify 7 of the highlighted errors and incorrectly identify 2 of the highlighted errors in Essay 3:

For your corrections in Essay 3 you receive _____
The participant whose essay you are editing will receive _____

4. If you correctly identify 4 of the highlighted errors and incorrectly identify 4 of the highlighted errors in Essay 4:

For your corrections in Essay 4 you receive _____
The participant whose essay you are editing will receive _____

5. If you correctly identify 4 of the highlighted errors and incorrectly identify 0 of the highlighted errors in Essay 5:

For your corrections in Essay 5 you receive _____
The participant whose essay you are editing will receive _____

6. If you correctly identify 0 of the highlighted errors and incorrectly identify 0 of the highlighted errors in Essay 6:

For your corrections in Essay 6 you receive _____
The participant whose essay you are editing will receive _____

7. If you correctly identify 9 of the highlighted errors and incorrectly identify 5 of the highlighted errors in Essay 7:

For your corrections in Essay 7 you receive _____
The participant whose essay you are editing will receive _____

8. If you correctly identify 1 of the highlighted errors and incorrectly identify 6 of the highlighted errors in Essay 8:

For your corrections in Essay 8 you receive _____
The participant whose essay you are editing will receive _____

9. If you correctly identify 3 of the highlighted errors and incorrectly identify 1 of the highlighted errors in Essay 9:

For your corrections in Essay 9 you receive _____
The participant whose essay you are editing will receive _____

10. If you correctly identify 0 of the highlighted errors and incorrectly identify 6 of the highlighted errors in Essay 10:

For your corrections in Essay 10 you receive _____
The participant whose essay you are editing will receive _____

11. Please calculate your total earnings from correcting these ten essays (questions #1-10)

Total Earnings _____

APPENDIX C:

Instructions

This is an experiment in decision-making. You will have an opportunity to earn a considerable amount of cash through your participation in this experiment. You will be required to complete the experiment individually. You are not allowed to communicate with any other participant at any point during the course of the experiment. If you complete the experiment, you will be paid for the decisions you make in the task. You may withdraw from the experiment at any time. If you withdraw from the experiment before its completion, you will only be paid for the portion of the experiment that you have completed. If you do remain in the experiment, then you should feel free to try to make as much money as you can. All responses and decisions will be anonymous and the only piece of experimental material that will contain your identity will be your receipt of payment at the end of the experiment.

Basic Overview

In this experiment, you are asked to proofread several short essays for spelling and typographical errors. At the end of the experiment you will be paid for your performance in this proofreading task.

Earnings Task:

Each person starts the experiment with $25. You will be given 10 essays to proofread. Each essay contains 10 spelling and typographical errors. For each error you correctly identify there is no change to your earnings, however for each error in the essays that you fail to correct you will have $0.25 taken away from your $25.00. You will not be penalized for changes made to portions of the text that did not originally contain an error.

You will have a total of 50 minutes to complete the proofreading task. At the end of the 50-minute period, the instructor will collect the essays to check for correctness and calculate your earnings.
For example:

- If in Essay 1, if you correct 5 out of 10 errors, $1.25 was taken away from your $25.00.
- If in Essay 2 if you correct 0 out of 10 errors, $2.50 was taken away from your $25.00.
- A total of $3.75 would be taken away from your $25.00 in this example.

Note: total earnings will be calculated over all 10 essays.

Future Earnings:

In future sessions of this experiment, other participants will proofread your corrected essays for the spelling and typographical errors that you were unable to correct, these participants will not know whose essay they are proofreading and you will not know who proofread your essay. This will allow you to earn some of your money back. For each error that they find that you failed to correct, we will award you $0.15.
For instance:

- If you missed 10 errors and the participants in the future sessions found 2 of them, you will receive $0.30.
- For this reason it is important that we have an email address to contact you so that you collect these future earnings.

CHAPTER III: OPTIMAL BLENDED PAYMENT STRUCTURES

ELLEN P GREEN
VIRGINIA TECH
DEPARTMENT OF ECONOMICS

1. INTRODUCTION

It is often the case that economist are unable to study the entirety of an economic problem due to time and budget constraints. Therefore they are forced to select between their treatments/hypotheses those of which they believe will provide the most interesting results. However, by methodologically selecting the treatments researchers can potentially predict their unobserved treatments'/hypotheses' outcomes through the use of counterfactual simulations. Given variation in the available economic data we can potentially predict unobserved outcomes by using counterfactual simulations to construct and predict the outcomes of the unobserved economic events from the available economic data. In this way, this methodology provides assistance to economic theory in modeling economic events when other constraints stand in our way. Through the use of deduction and data from a previously conducted experimental study, I investigate the impacts of blended payment mechanisms on agent behavior in the dual-principal agent relationship. Namely, I study the payment mechanisms that blend the incentives from the Salary, Piece Rate, Flat Rate and Bonus payment mechanism to find the optimal combination for a profit-maximizing firm.

In a previously conducted study by Green (2011), data reflecting the relationship between unblended payment mechanisms and agent behavior was collected in a laboratory experiment. The study shows the changes in agent behavior caused by the Flat Rate, Flat Rate with Bonus, Piece Rate, Piece Rate with Bonus, and Salary payment mechanisms. Green’s study is novel in that it is the first experiment that compares the effect of other-regarding preferences on within the dual-principal agent relationship.

To construct the effect of blended payment mechanisms on human behavior in the dual-principal agent relationship, I employ data collected from the experimental study conducted by Green (2011). From this I develop a model to fit the behavior induced by the payment mechanisms as seen in her experiment. The model specifically addresses the employer’s profit function, the agent’s profit function, and the client’s utility function. In the dual-principal agent relationship the employer is the contract designer; therefore I optimize for the employer’s profit through blended payment mechanisms.

The dual-principal agent relationship involves two principals interacting with a single agent. Here the upstream principal (the employer) in charge of contract design does not receive the services provided by the agent and the downstream principal (the client), receiving the services, does not directly pay the agent. This relationship presents itself most commonly in the Health Care industry as the relationship among the Physician (the agent) the Health Care provider (the employer) and the patient (the client). Of particular interest to this study is a dual-principal agent relationship where the optimal number of services provided to the client has an interior solution; meaning that services provided in excess of this quantity harm the client and results in a lower utility level. Also, if the number of services provided is less than this quantity the client’s utility will not be at its maximum level. In the health care industry example, patients require a specific optimal number of services, albeit unknown, from their physician in order to have ideal health. In the case that the physician over-provides services, the patient is negatively affected both monetarily and often times physically. On the other hand, fewer services than necessary also results in lower than optimal health.

The paper continues as follows. In the next section I provide a literature review, following this in section III is a description of the model employed in the optimization problem; next I provide a description of the data used in the model specifically detailing the experiment ran in Green (2011), with a description of the optimization process. In the final section I provide the conclusions from the optimization process and potential future projects.

2. Literature Review

The benchmark relationship between a principal and an agent and the problem of asymmetric information has resulted in several investigations exploring the effect of numerous variables (e.g. the number of parties, frequency of transaction, hidden information, etc.). The objective is to develop a contract that aligns the goals of the principal and the agent through both monetary and nonmonetary incentives (Pratt and Zeckhauser 1985; Eisenhardt 1989; Sappington 1991). These investigations have advanced the efficiency of contract design and their implications beyond this two-player relationship in primary contexts.

Of the many different contract designs, the most frequently discussed/complained about in economic literature are also those that are most frequently employed in the dual-principal agent relationship. Detailed explanations of the incentives of each the Flat Rate, the Piece Rate, and the Salary payment mechanisms and the failed incentives that they provide to the agent are often discussed in the dual-principal agent relationship in the health care industry (Prendergast 1999; MacNeil 1978; Rodwin 1993). The Flat Rate payment mechanism, which pays a flat rate for each patient on a physician's panel, encourages physicians to serve a large number of patients, however to under-provide services on a patient-to-patient basis. Additionally, it discourages physicians from taking on patients in their panel that will require more effort because they are less profitable for the physician then the patients that require less effort (Newhouse 1996). The Piece Rate mechanism pays a flat rate for each service provided by the physician, therefore rewarding physicians for over-providing services to the client is both cost inefficient and potentially detrimental to the well being of the client. Lastly, the Salary payment mechanism provides no link between the services provided and payment of the physician, encouraging laziness on the part of the physician (Hellinger 1996). While each of these payment mechanisms has their flaws, they also present benefits in the dual-principal agent relationship, as studied here with an interior solution.

In his paper on physician payment mechanisms, Robinson (2001) defines four features in which the contract designer (the employer) should specifically address in the dual-principal agent relationship with an interior solution. [I] Physician productivity and service. [II] Risk acceptance. [III] Efficiency and appropriate scope of practice. And [IV] Cooperation and evidence based-medicine. The most relevant to this study are Features [1], [II], and [III].[17] Feature [I] states that physicians should be

[17] This study focuses on the one time interaction between a single agent and their clients and therefore does not allow for cooperation amongst partners.

encouraged through their contract to provide as many services as necessary for a patient to be at their ideal health, Piece Rate encourages this by rewarding physicians for every service provided. Feature [II] states that contracts should acknowledge the additional risk physicians incur for taking on sicker patients by rewarding them for this burden, again this is encouraged through the Piece Rate payment mechanism, however discouraged in the Flat Rate payment mechanism. Feature [III], emphasizes efficiency in treatment. The Flat Rate payment mechanism promotes efficiency through risk sharing, which encourages the physician to invest their time in determining the most efficient treatment the first time around. While Feature [I] encourages the physician to do what is necessary, Feature [III] limits these treatments to those that are efficient.

While each of these payment mechanisms independently encourages at least one of these features, they do not satisfy all three. Therefore, we can hypothesize that a blended payment mechanism, which combines the incentives presented by the Piece Rate with the incentives of the Flat Rate, would result in a more cost-efficient and effective physician (Newhouse 1996). By this reasoning a payment mechanism constructed in this way encourages cost-effectiveness through the Flat Rate payment mechanism and encourage necessary treatment through the Piece Rate payment mechanism.

Green (2011) explores the effect of unblended payment mechanisms (Flat Rate, Piece Rate, and Salary) in a dual-principal agent relationship and finds that agency theory accurately represents the relationship between payment mechanism and quantity of services. The dual-principal agent relationship is presented as a one-time transaction between the client, employers, and agents with hidden information with respect to the effort levels of the agents. Within Green's investigation, she is able to directly quantify the quality of services provided by each of the payment mechanisms, which is extremely difficult to do in the field. Green finds a significant difference in the quality of services between payment mechanisms. Her study showed that in a comparison between unblended payment mechanisms, the Salary and the Flat Rate payment mechanisms provide the highest quality of services. These two payment mechanisms were able to outperform the Piece Rate payment mechanism, because Piece Rate agents were given a monetary incentive to over-provide services to the client the likelihood of a service being harmful is greater. However, in terms of cost effectiveness, the Piece Rate mechanism provided a better outcome from the employer's perspective, in that it has the highest profit margin. In the optimization presented below, I take these results and blend the above payment mechanisms in order to find a payment mechanism that provides an improved outcome and comes closer to addressing the three relevant features as described by Robinson (2001).

Table I: Treatments

Salary	Receives $25.00 at the end of the experiment
Flat Rate	Receives $2.50 for each client assisted
Flat Rate with Bonus	Receives $2.50 for each client assisted, plus a bonus $1.00, if conducted at least 9 of the quota errors
Piece Rate	Receives $0.20 for each highlighted section of the text changed
Piece Rate with Bonus	Receives $0.20 for each highlighted section of the text changed, plus a bonus $1.00, if conducted at least 9 of the quota errors
Socialization (Flat Rate Base)	Receives $2.50 for each essay edited and every period a table reflecting the other subject's behavior was displayed in the computer interface

3. METHODOLOGY

A. DATA

The data used in this study is drawn from Green (2011). Using experimental methodology Green investigated the different incentive structures of the Flat Rate, Piece Rate, Piece Rate with Bonus, Flat Rate with Bonus, and Salary payment mechanisms and agent behavior (as described in Table I). To determine their impact on human behavior of these payment mechanisms Green employed the "Dual-Principal Agent game" in a laboratory experiment with student subjects. The "Dual-Principal Agent Game" is a two-part game that assigns the role of the employer to the experimentalist, the role of the client to a one set of students and the role of the agents to another set of students. The clients in the experiment seek proofreading assistance for short essays in which there are spelling and grammatical errors. The agent's provide proofreading services to the clients' essays and are paid with respect to the services provided dependent on the payment mechanism the employer has set in place.

As discussed in the literature review, each of the payment mechanisms studied by Green are theoretically linked with the poorest of outcomes in the dual-principal agent relationship with an interior solution. The prospective payment mechanisms encourage agents to under provide services. On the other hand, the retrospective payment mechanisms, provide incentives for the agent to over-provide services. Green's findings conclude that while the relative frequency of services between payment mechanisms was accurately predicted by agency theory. However, the quality provided, which is more important, does not directly correlate to the frequency of services. For instance, although the Salary and the Flat Rate payment mechanisms did incentivize the agent to provide the fewest number of services they did not provide the minimum necessary to maximize their profit and also the client's product was of a higher quality than if they had been paid otherwise. Here, I combine these incentives from the prospective and retrospective payment mechanisms to form an optimal payment mechanism based on Green's findings and the following dual-principal agent model.

B. MODEL

Here agents have increasing effort costs and there is an interior optimum service level for the client (the downstream principal). In the model, agent j's effort on behalf of a client i as the number of services, s_i, where client i's utility is optimized at some value s_i^*; either more or fewer services than this are undesirable. However, s_i^* is not known with certainty, and the level of uncertainty varies across actors. Clients have the greatest uncertainty, agents face the least uncertainty due to their expertise and direct information pertaining to the client, and employers have more expertise than the clients but lack the agents' direct information, and thus they have an intermediate level of uncertainty about s_i^*.

In the model there are $i = 1...M$ clients and $j = 1...N$ agents where all agents report to a single employer. The employer must decide the optimal payment mechanism for the agent:

$$w_j(c^J, s^J, s^{J*}) = R_F \cdot c^J + R_P \cdot s^J + R_S + Z(s^{J*}), \tag{1}$$

where R_F denotes the Flat Rate wage vector, R_P is the Piece Rate wage vector, R_S is the Salary wage and $Z(s^{J*})$ is the bonus mechanism employed. In Green (2011) the bonus was earned when a pre-specified number of "quota" services were provided to the client. I describe c^J as the vector of clients served by agent j, s^J as the vector of services rendered by agent j and s^{J*} as the optimal number of services rendered by agent j. These vectors will be used to calculate the wage earned by the agent for each of the payment mechanisms employed.

The employer faces the following profit function:

$$\pi = \sum_i F_i(s_{ij}, s_i^*) - \sum_i G_i(s_{ij}, s_i^*) - \sum_j w_j(c^J, s^J, s^{J*}) - H. \tag{2}$$

$F_i(s_{ij}, s_i^*)$ describes the fees paid to the employer by the client for services rendered, which is a function of any agent j's services rendered to client i, s_{ij}, and the ideal number of services rendered to client i, s_i^*. $G_i(s_{ij}, s_i^*)$ is the cost or benefit incurred by the employer when any agent j provides services to services to client i, this could. For example, the cost of malpractice lawsuits as a result of over-providing or under-providing services to a client or the benefit of repeat clients in the long run. H represents other costs for the employer. Facing this profit function, the employer will choose the values of R_F, R_P, and $Z(\bullet)$ to maximize their profit. The employer can estimate the optimal service level, with some degree of uncertainty, and can use the estimate to design a bonus mechanism to encourage optimal behavior from the agent in the form of service quotas.

The client's utility in our model is described by the following function:

$$U_i = B_i(s_{ij}, s_i^*) - F_i(s_{ij}, s_i^*). \tag{3}$$

The utility of client i is calculated by the benefits of the services rendered, $B_i(s_{ij}, s_i^*)$, less the cost due to services to the client, $F_i(s_{ij}, s_i^*)$. $B_i(s_{ij}, s_i^*)$ reflects the change in utility to the client for being over-provided or under-provided services by agent j. Due to imperfect information *ex ante*, the client is unaware whether or not to decline services, and they are unable to accurately estimate s_i^*. In our model, the clients have minimal decision-making ability over the number of services rendered to them. Therefore I want to investigate situations in which the client is better off hiring an agent than not.

The agent's utility is defined by the following equation:

$$U_j = w_j(c^J, s^J, s^{J*}) - \sum_{i \in c^J} e_j(c^J, s^J, s^{J*}) + A_j(c^J, s^J, s^{J*}), \tag{4}$$

dependent on the agent's payment mechanism $w_j(c^J, s^J, s^{J*})$, the agent's effort function $e_j(c^J, s^J, s^{J*})$, and the agent's other-regarding behavior, $A_j(c^J, s^J, s^{J*})$. $A_j(c^J, s^J, s^{J*})$ describes the agent's other-regarding benefits with respect to client i, as described by Forsythe et al. (1994) and others. For example, if the agent is altruistic, this function could be described by:

$$A_j(c^J, s^J, s^{J*}) = m_i(B_i(s_{ij}, s_i^*)). \tag{5}$$

In this case the agent's utility increases when the client benefits from their services. The altruism multiplier is unique for each agent and I assume that $0 \le m_i \le 1$. The multiplier m_i reflects the influence of the client's net benefit on the agent's utility; the greater the value of m_i the more influence the client's net benefit has on the agent's utility. Depending on the magnitude of m_i, not addressing an agent's other-regarding behavior when constructing a utility function will lessen a theory's predictive power.

C. OPTIMIZATION AND METHODOLOGY

I. OPTIMIZATION

We maximize the following:

$$Max_X\ \pi = \sum_i F_{ij}(s_{ij}, s_i^*) - \sum_i G_i(s_{ij}, s_i^*) - \sum_J w_J(c^J, s^J, s^*) - H \qquad (6)$$

Where X is the vector of payment structures:

$$X = [R_S, R_F, Z(s^{J*}), R_P\ ] \qquad (7)$$

This optimization accounts for the utility of the client through their influence on the employer's profit as shown in $G_i(s_{ij}, s_i^*)$. Here I assume that $G_i(s_{ij}, s_i^*)$ represents the cost to the employer due to over providing or under providing services to the client. Therefore, let $G_i(s_{ij}, s_i^*)$ be defined as:

$$\sum_i G_i(s_{ij}, s_i^*) = f \sum_i F_{ij}(s_{ij}, s_i^*) - g \sum_i B_i(s_{ij}, s_i^*) \qquad (8)$$

Where g is the multiplier that represents the impact on the employer's profit caused by services performed for client i and f is the impact caused from the fees client i pays. Plugging this into the maximization problem above, we now have:

$$Max_X\ \pi = g \sum_i B_i(s_{ij}, s_i^*) + (1 - f) \sum_i F_{ij}(s_{ij}, s_i^*) - \sum_J w_J(c^J, s^J, s^*) - H \qquad (9)$$

For the purposes of this study, I assume that the fees paid by the clients directly equal the cost paid by the employer.

$$\sum_i F_{ij}(s_{ij}, s_i^*) = \sum_J w_J(c^J, s^J, s^*) + H \qquad (10)$$

Plugging this into our maximization problem we have:

$$Max_X\ \pi = g \sum_i B_i(s_{ij}, s_i^*) - f\{\sum_J w_J(c^J, s^J, s^*) + H\} \qquad (11)$$

Here the goal of our employer is to maximize in terms of the beneficial actions provided to their clients, in order to increase future transactions between the firm and the client.

II. METHODOLOGY

I assume that dataset reflects rational agents in a dual-principal agent relationship with other-regarding preferences. Therefore, the results to characterize and maximize an employer's profit, π, given the payment mechanism X in the above maximization problem. Here, the benefits to the client are defined as the net impact on their income caused by the services provided by the agent. In the "Dual-Principal Agent game" the clients received \$0.15 for necessary services, u, rendered and lost \$0.05 for unnecessary services, e:

$$B_{ij}\left(s_{ij}, s_i^*\right) = .15\left(u_J\right) - .05\left(e_J\right) \tag{12}$$

To determine the probability that the agent's services are beneficial to the client, we must determine the number of actions the agent performs given the payment mechanism. The observed agent's actions where fit using Ordinary Least squares regression on the payment mechanism as regressed on the $\log\left(s_i^J + 1\right)$:

$$\log\left(s_i^J + 1\right) = \beta_0 + \beta_1 X_i + \varepsilon_i. \tag{13}$$

From this regression, the fitted values were used to calculate the errors and the variance in actions. The justification for using to a log normal is that the data on services provided by the agents in Green (2011) showed a distribution similar to the log-normal density. The parameters taken from the original regression were then used to construct the expected value of the agent actions as a result of the new payment mechanisms. Let X_i^* define the new payment structures. Thus we can use the fitted values to calculate a new

$\sim \ln N(\mu,\ \sigma^2)$, where $\mu = \hat{\beta}_0 + \hat{\beta}_1 X_i^*$, $\sigma = var(\varepsilon_i)$. And thus have that the expected services provided by the agent as[18]:

$$E(\log(s_i^J + 1)) = e^{(\hat{\beta}_0 + \hat{\beta}_1 X_i^* + \frac{1}{2}\sigma^2)} \tag{14}$$

To determine the probability of the action taken to be beneficial to the client, I used the logit model on a dummy variable given the value of 1 if the service provided was beneficial and 0 if it was harmful to the client was created. Then running the logit regression with the dummy variable as the dependent variable and the payment vector, X, as the regressor I was able to determine the probability of the agent conducting beneficial treatments, u_J: $Prob(u_J|s_i^J, X)$.

[18] See Appendix A.II for the regression results.

Table 2: Optimal Payment Mechanism

Payment Structure	[0, 0.625, 0, 0.1113]	[25, 0, 0, 0]	[0, 2.5, 0, 0]	[0, 2.5, 1, 0]	[0, 0, 0, .20]	[0, 0, 1, .20]
Average Wage	$11.20	$25.00	$23.95	$23.06	$12.27	$15.64
Average Utility	$4.75	$5.65	$4.81	$54.15	41.333	0.6789
Beneficial Actions	35.5408	38.5204	32.4145	37.4822	36.0007	22.461
Harmful Actions	11.5927	2.48035	1.063	4.1432	25.336	53.8041
Profit	$9.71	-$0.12	-$2.78	$0.77	$5.92	-$12.65

*Payments given by: [Salary, Flat Rate, Bonus, Piece Rate]
**G=4.4, F=1

In calculating the wage of the agent we need both the probability of receiving a bonus and the probability of providing services for a client. The bonus structure used in Green's experiment was representative of a quota bonus as employed in the health care industry. If agents corrected at least 9 of the specific type of error in the essays, they were rewarded with a bonus. To determine the probability of an agent receiving a bonus, I ran the logit on a dummy variable for receiving a bonus on payment mechanism: $Prob(Z(s^{J*})|X)$.[19] The probability of an agent providing services was determined by using a logit regression on each client, regressing a dummy variable for client selection on the payment mechanisms.[20] These provided me with the probability that an agent would provide services for each individual client: $Prob(c^J|X)$.[21] To calculate the total number of clients served I summed the probability over all clients.

To find the optimal payment mechanism I used the interior point algorithm, with the barrier function defined by the maximum number of services that an agent can provide to their client panel (Karmarkar 1984; Byrd 1999). The maximum number of services the subjects could perform in the experimental study by Green to 135 over all clients, the following equation bounded the number of potential services by the constraint $X^*\hat{\beta} \leq 4.8248$. Where $\hat{\beta}$ is a vector of coefficients from the ordinary least squares

[19] See Appendix A.II for regression results.
[20] See Appendix A.III for the regression results
[21] See Appendix A.IV for regression results

regression as described in equation (13), $\hat{\beta} = [3.4385, 3.4303637, 3.6482, 6.4579]'$. Additionally, the maximization function was bounded such that none of the values within the payment structure vector could be less than zero: $X^* \geq [0, 0, 0, 0]$. In Green's study, the payment mechanisms were framed as a reward incentive, rather than a punishment. Therefore, based on the results of Fehr et. al (2002), which showed individual's responded with increased effort when incentives were framed as positive rewards rather than punishments, it would be inappropriate to allow for punishments in this optimization. Another requirement in the optimization problem was that at least one value in the vector must have been greater than \$0.05. The reasoning behind this constraint is that in previously conducted experiments, it has been shown that the minimum necessary wage to encourage variation in behavior due to monetary incentives was \$0.05 (Smith 1976).

Table 3: Payment Mechanism

Payment Structure	**[10, 0, 0, .1]**	**[10, 1, 0, 0]**	**[15,0,0,.1]**	**[1, 0.625, 0, 0.1113]**	**[0, 2.0, 1, .2]**	**[0, 2.0, 0, .2]**
Average Wage	\$14.91	\$19.31	\$20.12	\$12.13	\$33.67	\$31.35
Average Utility	\$5.28	\$4.91	\$6.21	\$4.93	\$12.27	\$8.51
Beneficial Actions	38.6764	33.6466	43.8238	36.4637	63.3171	57.8605
Harmful Actions	10.4688	2.661	7.3565	10.8098	12.9481	3.4762
Profit	\$8.31	\$2.31	\$7.19	\$9.56	\$5.27	\$6.08

*Payments given by: [Salary, Flat Rate, Bonus, Piece Rate]
**G=4.4, F=1

In this optimization, I assume that the outcome on the client directly impacts the profit of the employer on a 1:1 scale, therefore let $f = 1$ and increased the value of g from 0 to determine the lowest possible value at which it was profitable for the employer to hire the agents (i.e. the optimal payment mechanism is greater than 0 for all inputs). Note that if both f and g were assumed to be less than this value for g, there would be no incentive for the employer to hire the agents; both principals would be better off if they had no contracts with the agents.[22] After running the optimization as described above, the lowest value at which there was economic profit in hiring the agents was at $g = 4.4$, at any value less than 4.4 the optimal payment mechanism was $X = [0, 0, 0, 0]$.

Using this methodology, I found that the optimal payment mechanism is a blend between the Flat Rate with Bonus and Salary payment mechanism where $X^* = [0, .625, 0, .113]$. Employing this payment mechanism in the model above results in a profit to the employer of $11.59, as shown in **Table 2**. The average net benefit for all 10 clients was $4.75. The rate of beneficial services conducted by the agent was 96% under this payment mechanism. The average wage paid to the agent was $11.20. Table 2 also lists the expected outcome of the unblended payment mechanisms explore in Green (2011). Out of the tested payment mechanism, only two where profitable to the employer at $g = 4.4$. These were the Flat Rate with Bonus and Piece Rate payment mechanisms.

When determining an optimal payment mechanism, one must address both the cost and benefits. By increasing the wage, we can increase the benefits received by the clients, however, this is not cost efficient. Table 3 provides alternative payment structures as a comparison to the optimal payment mechanism. In an instance were the benefits are larger than in the optimal payment structure, X^*, we can pay our agents a combination that includes Salary, Flat Rate, and Piece Rate: $X = [1, .625, 0, .113]$. Here the utility received by the clients is $4.93, which is greater than the utility provided by X^*, however the wage is increased because we are now including the Salary wage at $1.00. When $X = [10, 0, 0, .1]$ a combination of the incentives from the Salary payment mechanism (at

[22] The author acknowledges that according to this logic, Green (2011) should not have completed the experiment, as the experimenters were defined as the employer.

$10) and Piece Rate payment mechanism (at $0.10) the employer would profit at $8.31. The average utility of all clients would higher at $5.28 and however the Wage would be more costly at $14.91. While these other blended payment mechanisms provided higher utility, they also cost more to implement. Therefore, according to this maximization problem they are not the optimal.

6. CONCLUSIONS

Here I have simulated an agent's payment mechanism based on existing experimental results, in terms of employer's profit. In turn, this also investigates the optimal payment mechanisms in terms of the client impact and cost to the employer. For example, the maximization function is in terms of benefits to the client and benefits to the employer. In this way, I minimize the wage paid out by the employer and maximize the benefits received by the client. From this optimization we see that the combination of Flat Rate and Piece Rate payment mechanisms provide the optimal payment structure within this dual-principal agent relationship. While each of these payment mechanisms on their own provides flaws, such as encouraging overtreatment, their combined effects result in an improved outcome for both principals as optimized here.

As described by Robinson (2001) there are four key features in contract design for a dual-principal agent problem with an interior solution. The Piece Rate's ability to encourage agents to provide as many services as possible (Features [I] and [II]) results in the services being over provided. In the dual-principal agent problem investigated here, there is an interior solution to the optimal number of services provided and hence, the Piece Rate is shown to over provide and be detrimental to the overall well being of the client. Conversely, the Flat Rate payment mechanism encourages the agent to be cost effective (Feature [III]), however it also does not adjust agent payment to appropriate reflect providing services to the more time consuming clients. Here we see, as predicted, that their offsetting advantages provide an efficient payment mechanism in terms of employer profit and the overall outcome of a dual-principal agent relationship.

7. References:

Byrd, R.H., Mary E. Hribar, and Jorge Nocedal, "An Interior Point Algorithm for Large-Scale Nonlinear Programming," *SIAM Journal on Optimization*, Vol 9, No. 4, pp. 877–900, 1999.

Eisenhardt, K.M. 1989. "Agency Theory: An Assessment and Review." *Academy of Management Review.* 14 (4): 57-74

Fehr, E., U. Fischbacher, et al. (2002). "Strong reciprocity, human cooperation, and the enforcement of social norms." *Human Nature-an Interdisciplinary Biosocial Perspective* 13(1): 1-25.

Forsythe, R., J. L. Horowitz, et al. (1994). "Fairness in Simple Bargaining Experiments." Games and Economic Behavior **6**(3): 347-369.

Green E. "Payment Mechanisms in a Dual-Principal Agent Setting: An Experimental Study." *Working Paper.* 2011

Hellinger, F.J. 1996. "The Impact of Financial Incentives on Physician Behavior in Managed Care Plans: A Review of the Evidence." *Medical Care Research and Review.* 53(3): 294-314.

Karmarkar, N. 1984. "A New Polynomial-time Algorithm for linear programming." *Combinatorica*, 4(1984) p.373-395.

MacNeil, I.R., 1978. "Contracts: Adjustments of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law." *Northwestern University Law Review* 72:854-906.

N. Hansen, S.D. Muller, P. Koumoutsakos, *Evolutionary Computation*, **11**(1), 1–18, (2003).

Newhouse, J.P. 1996. "Reimbursing Health Plans and Helath Providers: Efficiency in Production versus Selection." *Journal of Economic Literature* 34:1236-63.

Prendergast, C. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37(1): 7-63.

Pratt, J.W. and R.J. Zeckhauser. 1985."Principals and Agents: The Structure of Business." Cambridge, Mass.: *Harvard Univserity Press*.

Robinson, J. C. (2001). "Theory and practice in the design of physician payment incentives." *Milbank Quarterly* **79**(2): 149.

Rodwin, M.A. 1993. “Medicine, Money and Roals: Physicians’ Conflicts of Interest.” *New York: Oxford University Press.*

Sappington, D.E.M, 1991. “Incentives in Principal-Agent Relationships. *Journal of Economic Perspectives* 5:45-66

Smith, V.L., 1976, Experimental economics: Induced value theory, The American Economic Review, 66, 274- 279.

8. APPENDIX:

TABLE A.I:

Ordinary Least Squares regression with Log(Actions+1) as the dependent variable	
Salary	-0.0222***
	(-3.45)
Flat Rate	-0.303***
	(-6.23)
Bonus	0.218*
	(2.24)
Piece Rate	0
	(.)
Constant	4.036***
	(46.51)
N	106

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

TABLE A.II

Logit regression with Beneficial Dummy as the dependent variable	
Salary	0.0180***
	(19.32)
Flat Rate	0.223***
	(33.42)
Bonus	-0.231***
	(-18.23)
Piece Rate	0
	(.)
Constant	0.546***
	(48.75)
N	5252

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

TABLE A.III:

Logit regression with Bonus Dummy as the dependent variable	
Salary	0
	(.)
Flat Rate	-0.536*
	(-2.07)
Bonus	1.222
	-1.88
Piece Rate	0
	(.)
Constant	-1.696**
	(-2.99)
N	84

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

TABLE A.IV:

Client Selection by Payment Structure from Logit Regression

	Client 1	Client 2	Client 3	Client 4	Client 5	Client 6	Client 7	Client 8	Client 9	Client 10
Salary	0	0	0	0.00441	7.05E-19	0.000598	-0.00102	0.000101	-0.0155*	-0.0121
	(o)	(o)	(o)	(1.43)	(0.25)	(0.25)	(-0.26)	(0.02)	(-2.61)	(-1.84)
Flat Rate	0	0	0	0.0123	1.08E-17	-0.00866	0.000911	0.0141	0.0524	0.0961
	(o)	(o)	(o)	(0.66)	(o)	(-.59)	(0.04)	(0.44)	(1.46)	(2.40)
Bonus	0	0	0	0.0253	-0.0476	-0.0248	-0.143	-0.0697	-0.113	-0.06
	(o)	(o)	(o)	(0.54)	(-1.60)	(-0.68)	(-2.37)	(-0.88)	(-1.26)	(-.06)
Piece Rate	0	0	0	0	0	0	0	0	0	0
	(o)	(o)	(o)	(o)	(o)	(o)	(o)	(o)	(o)	(o)
Constant	1	1	1	.912***	1***	.988***	0.975***	0.862***	0.809***	.606***
	(o)	(o)	(o)	(22.81)	(37.46)	(30.15)	(18.09)	(12.15)	(10.10)	(6.79)
N	106	106	106	106	106	106	106	106	106	106

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

CHAPTER IV: RESPONSIBILITY AND INTENTIONALITY IN ECONOMIC GAMES

BY:

ELLEN GREEN, SHERYL BALL, LANDIS ATKINSON, AND AIDA CRUZ-FOLGAR

VIRGINIA TECH

DEPARTMENT OF ECONOMICS

BLACKSBURG, VA 24061

1. INTRODUCTION:

The breadth of experimental economics literature has led us to reconsider traditional theory and address behavior displayed by humans in everyday situations. In particular, our concern lies in the results of an individual's intentions and responsibility effect on decision-making. In recent years, experimental economists have explored intentions effect on human behavior. Previously, intentions had been ignored in economic theory on decision-making; however, economic experiments have since shown the impact of intentions on the decision-making process (Blount, 1995). Falk, Fehr, and Fischbacher (2008) advanced decision-making theory by showing that when intentionality was removed from a monetary interaction, individuals had less intense reciprocal behavior, meaning they sanctioned and rewarded their counterparts less. On the other hand, when intentions were relevant, individuals were more likely to sanction (reward) negative (positive) behavior. Our experiment expands on Falk, Fehr, and Fischbacher (2008) by exploring how the assignment of responsibility within a decision-making process affected the outcome of a similar monetary interaction. In other words, how does taking responsibility for one's actions affect the responses in a monetary interaction between individuals?

The primary innovation of this paper is to show what happens when two anonymous individuals are given the opportunity to take direct responsibility for their actions, rather than having the decision-making process assigned to them by the experimenter as in past experiments. We define *direct responsibility* as when an individual is directly responsible for the outcome. Conversely, an individual is considered to have *indirect responsibility* for an outcome if they have chosen to assign the decision making task to another entity outside of the players within the game. The decision made by another entity had no impact on the entity's income; they were not invested in the outcome of their decision. Individuals in our experiment are able to have either indirect responsibility for the outcome or direct responsibility for the outcome. This is achieved by allowing the participants to decide whether they preferred to distribute a pre-specified amount of money via a random device or perform the distribution task themselves.

Selecting the random device is akin to removing responsibility from the decision-making process and having indirect responsibly for the outcome because they are no longer directly in charge of making the decision. Conversely selecting to distribute the wealth themselves relates to taking direct responsibility of the outcome. After assigning responsibilities for the actions, participants were also able to show their intentions through the distribution of the wealth, by acting selfishly (by taking from their counterpart) or behaving kindly (by contributing to their counterpart's wealth). The counterpart was then given the opportunity to show their reaction to the distribution of money by either sanctioning the original player (taking money away from them, with a cost to the punisher) or rewarding the original player (gifting the original player, with a cost to the gifter).

There are many occurrences in which the attribution of responsibility and the acceptance of responsibility are relevant. In particular, the application of the responsibility scenario occurs when a group of individuals are trying to make a decision that will affect all players involved. Consider the two following scenarios, two tourists, Ted and Tom, are out in Rome trying to determine whether to follow their guidebook or to hire a local tour guide to show them around town. In the first scenario, Ted decides to use the guidebook for the remainder of the day, he prefers to save money by not hiring local tour guide and prefers to take the chance that he can find all of the sites that they want to see and successfully entertain his Tom. On the other hand, Ted could have chosen to provide the tour himself knowing that he would not provide as good of a tour as an experienced tourist, but wanted to save the money. Independent of the reasoning, Ted is able ensure that the pair does not pay to have somebody give them an awful tour. If Ted had hired an individual to give them a tour, there is a chance that the tour guide would have an immense knowledge of the local attractions and provide a delightful tour allowing the two to see places that they might not have visited. By the same token, the hired tour guide could speak poor English and fail to visit the most important sites in the city. At the end of the day after either of these outcomes Tom reacts to how pleased he is with the day.

In another scenario, Tom says to the Ted, “you provide us with the tour.” He has faith in the tour guide ability of the Ted. Tom also believes that assigning the responsibility to Ted will encourage him work harder to entertain him (he is familiar with Charness (2009)). Or, Tom could elect to hire a tour guide with an unknown quality and play the odds that the hired tour guide would be better at providing an informative tour. Again, the potential outcomes as a result of Ted giving a tour or the tour guide providing a tour are possible. At the end of the day after either of these outcomes Tom reacts to how pleased he is with the day.

Both scenarios provide a first and second mover as well as the assignment of responsibility over the quality of tours provided. In the first scenario, the first mover is Ted, who gets to choose between a hired tour guide and a self-guided tour. The second move is either Ted’s quality of performance or the hired tour guide’s quality. The third move comes from Tom, when at the end of the day he reports his satisfaction with the outcome. In the second scenario, the first-mover is Tom who assigns the responsibility of the tour to either Ted or a hired tour guide. The second mover is Ted who is either following the lead of the tour guide or provided a self-guided tour for the pair and the quality is the outcome. Again, in the second scenario Ted has the opportunity to respond to the outcome in the third move.

What we want to understand with our experiment are the following, does it make a difference to Ted whether Tom provides a bad tour or the hired tour guide gives a bad tour? And does it matter who selected the tour guide or Tom to give the tour? A consequentialist would say that both of these cases should result in the same utility. However, our study showed that taking responsibility for one’s action increases response levels of their counterpart, in terms of rewards and sanctions. Meaning that in the above scenarios, there is a difference between who gives the tour and the satisfaction levels of

Tom. Additionally, we found that the person who made the decision to assign responsibility either directly or indirectly had a significant impact on the sanctioning and rewarding in the third move.

2. LITERATURE REVIEW:

A. THEORETICAL:

There are several degrees of intentionality as discussed in economic justice literature. There is the matter of whether an individual is responsible for the act as well as the consequences or if either of these aspects is not the responsibility of the individual it is considered unintentional (Bentham 1970). These features distinguish between two different contexts that theorist typically address in their papers. In this way individuals respond to the attribution or assignment of responsibility in different ways.

One mechanism individuals employ to disperse decision-making responsibility from them is a random devise (e.g. a flip of a coin). Psychologists explain this decision-making process through attribution theory, which proposes that individuals assign responsibility of an unfavorable outcome if the intentions are clear (Heider, 1958; Kelley 1967, 1973; Ross and Fletcher, 1985). And for this reason, intentions impact the human decision-making process. Traditional consequentialist economic theory has ignored the impact of intentions and responsibility, however in recent years they have begun to incorporate these variables into theoretical models. Dufwenberg and Kirchsteiger (2004) and Rabin (1993) both address the intentions of an individual in their models, Dufwenberg and Kirchsteiger in a sequential model and Rabin in a one time model. While, both describe reciprocal behavior on intentions in their investigations on decision-making to be incomplete, because they fail to incorporate the effect of the distribution of intentions on the behavioral outcome. Sebald (2009) directly incorporates the attribution of responsibly, into a game theoretical model of human behavior. In this way, he is able to successfully predict the outcomes of laboratory experiments on intentions and its influence on the human-decision making process. In particular, those experiments in which chance moves are involved. These experiments are described in the following section.

B. EXPERIMENTAL:

I. INTENTIONS:

Recent work in economics has found that reciprocal behavior depends not only on an individual's own payoff and the distribution of payoffs across players, but also on the intentionality of others. Here we explore the relevant influences on decision-making and more appropriately, model individual behavior with respect to the intentionality signaled by the players as well as those papers that address the attribution of responsibility.

Economic experiments have proven that other-regarding behavior has a significant effect on individual's decision-making (i.e. fairness and altruism) and in response, experiments have been developed to determine what other characteristics could alter the decision-making process. Blount (1995) first introduced the idea of examining the role of intent to the ultimatum game and tested the hypothesis that the intent of an individual would change the outcome of the game. Blount removes intentions by assigning the decision-making task in the first stage of an ultimatum game to an uninterested third party or a computer, which randomly determines the allocation. The results experiment show that the intention behind the proposal directly influences the acceptance levels of the respondent. Individuals assigned to the treatment without intentions were more likely to accept offers absolutely and were more likely to accept lower offers than those who participated in the third party treatment. Additionally, when compared to the results of a traditional ultimatum game, both of Blount's treatments resulted in higher acceptance levels. These results point to the conclusion that the intentionality of the proposer has a significant effect on the outcome of the game.

Falk, Fehr, and Fischbacher (2008) examined the impact of intentions in their experiment by comparing the behavior of subjects in the 'moonlighting game' with intention and without. The treatments in the Falk, Fehr and Fischbacher study where similar to those in the Blount study, in that the subjects were paired with either human or computer counterparts. What the Falk, Fehr, and Fischbacher study adds to the intentions literature are their observations on the positive and negative impact intentions have on reciprocal behavior. The 'moonlighting game' allows the respondent to either reward or punish their counterpart for satisfactory (or unsatisfactory) behavior. Here, individuals were more likely to behave with self-regarding preferences and were less likely to reciprocate/reward positive behavior when the decision-making was conducted by a computer rather than their human counterpart. Fehr et al's findings disprove theories such as Dufwenberg and Kirchsteiger (2004) and Rabin (1993) as described above because they do not address the distributional effect of intentions. Furthermore, Sebald (2009) directly accounts for the impact of distribution intentions and thereby is able to model the behavior seen in these papers.

II. RESPONSIBILITY:

In addition to the role of intentions on the decision-making process we explore the differences in the attribution of *indirect* responsibility and the *direct* responsibility. Fields in psychology, management, and education are dedicating to understanding the effect responsibility has on individual behavior and now its impact on economic decision-making is being developed. Charness has headed the work in experimental economics on responsibility in work environments (Charness 1998, 2000, 2009). His first paper on responsibility alleviation shows that a shift in responsibility to an external authority will reduce a counterpart's tendency towards honesty, loyalty, or generosity in the gift-exchange experiment. In another treatment, when responsibility is assigned to a random process, subjects were more generous seemingly indicating that this random process was more just than an external authority. Charness (2000) explores the effort effect over the responsibility of decision-making over wages and shows that humans have potential to act in 'pro-social' manners when assigned responsibility. In Charness (2009), the author continues to explore the effect of responsibility and its effect on risk-taking in a two-player game versus a one-player game. In this context, of the individuals that are affected by being assigned the responsibility of another person's welfare, about 90% of them chose to take on less risk when acting on behalf of another player.

Beyond these papers there have yet to be any studies that directly investigate the effect the attribution of responsibility has on individuals' decision-making process, which is the motivation behind our analysis. Here we allow individuals to directly or indirectly assign responsibility of the decision-making task and discover the differences in outcomes caused by this assignment.

3. EXPERIMENTAL DESIGN AND PROCEDURE

A. EXPERIMENTAL DESIGN:

Our experiment employs a modified version of the "moonlighting game" as introduced by Abbink, Irlenbusch, and Renner (2000) and follows closely to the "constituent game" as designed by Falk, Fehr, and Fischbacher (2008) (which is another interpretation of the "moonlighting game"). An important advantage of employing the "moonlighting game" over other experimental designs is that it allows the experimenter to study the effects of both positive and negative reciprocity, which is lacking in the traditional ultimatum game. In the "moonlighting game," the task is to divide and endowment of wealth between two players, Player A and Player B, who are each endowed with 12 points at the start of the game. Their task in the game is to allocate the endowment between them through a two-stage game. Player A is assigned the decision-making task in Stage 1 and Player B is assigned the decision-making task in Stage 2.

In Stage 1 Player A (or a device representing Player A) determines the amount of points that he will give or take from Player B. This amount can range from +6 and -6 where a positive number represents giving away points to Player A's counterpart Player B, and a negative number represents taking away points from Player A's counterpart Player B. The potential earnings of Player A and Player B are shown in Table I above. If Player A decides to give (gift) his counterpart Player B with X points, the points are tripled and Player B's payoff at the end of Stage 1, P_B, is 12+3X points from Player A (it will cost Player A X points to give to Player B in this way). If Player A decides to take away X points from his counterpart then Player A gets X points and Player B loses X points and his payoff at the end of Stage 1, P_B, is 12-X. Player A's payoff at the end of Stage 1, P_A, will be 12±X.

Table I:
Potential Outcomes in the" Moonlighting Game" in Stage 1

Player A's Decision (X)	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
Player A's Profit (P_A)	18	17	16	15	14	13	12	11	10	9	8	7	6
Player B's Profit (P_B)	6	7	8	9	10	11	12	15	18	21	24	27	30

*Note negative numbers represent taking from Counterpart

In Stage 2, Player B responds to the decision made in Stage 1 by Player A. Player B can either take away points (sanction) from Player A or give points (reward) to Player A. This amount can range between -6 and +18 depending on the Player B's balance after the decisions made in Stage 1 and all possible outcomes are presented in Table II. The payoffs in Stage 2 are calculated in the following manner, if Player B decides to give Y points to Player A she gives exactly Y points. If Player B decides to take away Y from Player A points then this costs her Y points, but reduces Player A's points by 3Y. In other words, reducing Player A's points is costly to Player B and is akin to paying to sanction Player A for the distribution of wealth. Using backward induction we find that the subgame perfect Nash equilibrium is for Player B to neither take or give money to Player A, because both of these actions decrease their payoff. Therefore, in the first stage, the best response for Player A is to take all that they can from Player B (i.e. 6 tokens). This is of course, assuming that both players act in a self-interested manner.

Table II:

Potential Outcomes in the" Moonlighting Game" in Stage 2: (Y)

Player B Decision	18	17	16	15	14	13	12	11	10	9	8	7	6
Player A's Profit	P_A+18	P_A+17	P_A+16	P_A+15	P_A+14	P_A+13	P_A+12	P_A+11	P_A+10	P_A+9	P_A+8	P_A+7	P_A+6
Player B's Profit	P_B-18	P_B-17	P_B-16	P_B-15	P_B-14	P_B-13	P_B-12	P_B-11	P_B-10	P_B-9	P_B-8	P_B-7	P_B-6

Player B Decision	5	4	3	2	1	0	-1	-2	-3	-4	-5	-6
Player A's Profit	P_A+5	P_A+4	P_A+3	P_A+2	P_A+1	P_A+0	P_A-3	P_A-6	P_A-9	P_A-12	P_A-15	P_A-18
Player B's Profit	P_B-5	P_B-4	P_B-3	P_B-2	P_B-1	P_B-0	P_B-1	P_B-2	P_B-3	P_B-4	P_B-5	P_B-6

*Note negative numbers represent taking from Count

Falk, Fehr, and Fischbacher (2008) add intentionality to the “moonlighting game” in the “constituent game.” Their modified version of the “moonlighting game” assigns the task in Stage 1 to either a random decision making device or a human counterpart. In this way, the investigators are exploring the impact of intentionality on the decision-making process of Player B by removing the decision-making task of Player A. In our investigation we add responsibility to the “constituent game”.

The key difference between the “constituent game” and the “responsibility game” is that players in the “responsibility game” are given the opportunity to assign responsibility to the decision-makers. This occurs in Stage 0, prior to the two stages in the traditional “moonlighting game”. In the “responsibility game”, in Stage 0 Player A or Player B decides whether Player A or a random device, as in the Falk, Fehr, and Fischbacher (2008) study, makes the decision in Stage 1. The random device employs the distribution of allocations listed in Table III. This distribution of allocations was taken from Abbink et al (2000) and represents those decisions made by participants in their study. This distribution was also used for the random device in the Falk, Fehr, and Fischbacher (2008) study. Column one represents a range of numbers that the device randomly generates, in our case a random number generator within the “z-tree” code. Column two represents the choice related to the generated numbers and the allocation by Player A. Column 3 shows the probability of each allocation of wealth by Player A in

Table III: Probabilty distribution of move in Stage 1

Realized number	A's move	Percent
0-6	-6	7
7-8	-5	2
9-15	-4	7
16-19	-3	4
20-21	-2	2
22-26	-1	5
27-39	0	13
40-46	1	7
47-55	2	9
56-62	3	7
63-73	4	11
74-75	5	2
76-99	6	24

Stage 1 if the device is chosen in Stage 0. In our experiment the players making the decision in Stage 0 were aware of the distribution that would be employed by the random device. This aspect is key to our analysis of the responses from the players. Providing this information to the players allows the players to be held fully accountable for their decision in Stage 0. In this way there is no doubt what could potentially happen if the random device was selected without any knowledge of the potential distributions by the device.

On the other hand, if responsibility is not a factor in the decision-making process, allowing the participant to see the probability distribution if the random device is employed would results in indifference between having their counterpart determine the distribution of wealth versus using a random device. As the player's expectations from a human decision-maker should be the same as the device. In this way, the “responsibility game” studies the effects of taking responsibility of one’s actions as well as the effect of intentionality (determined in Stage 1) after the attribution of responsibility (determined in

Stage 0). Player A's intentionality signal can range between kind and selfish in Stage 1 and the response of Player B in Stage 2 can result in sanctions or rewards. All ranges are important to observe proper analysis of the interaction.

As described above, the focus of the experiment is to determine how individual behavior changes with respect to assignment of responsibility. In the "responsibility game" one participant in the interaction was asked to make a decision regarding the mechanism for distributing wealth amongst the pair. In this way, we had two treatments, which can be described as follows:

i. A-Choose Treatment
- Player A is assigned the task in Stage 0 to determine the distribution of their own endowment themselves or to allow a device randomly determine the allocation.
- Individuals who select to make the decision themselves are considered responsible.

ii. B-Choose Treatment
- Player B is assigned the task in Stage 0 to determine the distribution of Player A's endowment by Player A or to allow a device randomly determine the allocation.
- Individuals who assign the decision making task to Player A are assigning responsibly to Player A.

Allowing the decision in Stage 0 to be made by different Players creates two entirely different scenarios. As it has been shown in other experimental literature, the assignment of authority has a significant effect on the allocation of wealth (Hoffman and Spitzer 1985). Individuals assigned to positions feel justified to behave more selfishly and furthermore, are rewarded with higher payoffs from their counterparts. This assignment of authority does not depend on the methodology (e.g. coin flip vs. ranking of quiz results), but rather that there has been an assignment at all. In this way, by entitling the players with the task of making the decision in Stage 0, we anticipate that the players will have a change in behavior as found in the previously conducted studies. This entitlement can potentially result in increased response levels in both the positive and negative.

There are two levels of assignment in the B-Choose treatment. First, Player B is assigned the task of decision-maker in Stage 0. Secondly, they are given the task of assigning responsibility for the decision made in Stage 1. Perhaps giving a sense of entitlement to Player A and potentially affecting the behavior of Player A in the allocation of wealth.

B. PREDICTIONS:

Fairness, altruism and many other human characteristics are often involved in decision-making tasks. As seen in previous studies the intentions behind decisions can have both a detrimental and positive effect on the final utility of an individual. However, these studies do not focus on how the responsibilities change the outcome. In the previous experiments discussed above the participants were not given the opportunity to signal their acceptance of responsibility for the outcome and consequently the factor of responsibility has been removed from the interaction. We feel that signaling responsibility will have a significant impact on future utility theory and have designed the "responsibility game" to take this aspect of social interactions into account.

Within the experimental design of the "responsibility game," individuals assign direct responsibility by selecting the allocation to be divided by a human rather than to assign the decision to a random device. If the decision is assigned to the random device the person accepts indirect responsibility for the outcome. After this, in the case that the player is making the decision on their own accord, they are able to signal intentions by either giving or taking from their partner. On the other hand, if individuals select to have the decision made by a random number generator, direct responsibility is lost. Given this scenario, we make the following predictions:

Prediction 1: Those players who are given the opportunity to take responsibility for their actions and do, will be rewarded for their kind behavior and sanctioned for their selfish behavior.

Prediction 2: Those players who are given the opportunity to take responsibility for their action and choose not to, will be rewarded for their kind behavior and sanctioned for their selfish behavior.

We predict that the two outcomes will differ from the findings of Falk, Fehr, and Fischbacher (2008) with regards to the participant's reaction to the random device. In the Falk, Fehr, and Fischbacher paper, participants paired with a random decision-making device were less likely to punish and reward their counterparts. We predict that poor allocations made by the random device are still punishable because Player A is still indirectly responsible for the outcome. In our tourist example, this would be the scenario when Tom decided to provide the tour by himself, however did a poor job at showing Ted the sights. Because Tom chose to take direct responsibility for the tour and did poorly Ted will resent the outcome. In turn, this resentment would still be observed when Tom chooses to hire a tour guide that does a poor job at exploring Rome because the decision was in Tom's hands to hire a tour guide. Therefore, he is indirectly responsible for the outcome. In both cases, Tom is responsible for the poor tour, however he has greater direct responsibilities in the case that he provides the tour himself. This difference in responsibility should present itself when Ted displays his dissatisfaction with the outcome, meaning he will be more upset at Tom for giving the poor tour himself over the tour guide's lack of skill.

These behaviors can also be explained by two different human motivations. First, by assigning responsibility of the decision-making task to a random device, the player is signaling to their counterpart that they are not willing to take responsibility for their actions. Humans are encouraged to take responsibility for their actions from a very young age, if children fail to provide satisfactory grades they are typically at the very least verbally sanctioned and conversely rewarded for good grades. Second, children who complain about poor outcomes being due to the teacher or other circumstances not under their control will (should) still be sanctioned for poor outcomes and rewarded for positive outcomes. Humans are motivated by the desire to create integrity in a situation, by assigning responsibility to actions.

Furthermore, we explore the responses of individuals when they are given the opportunity to assign the decision-making task in Stage 1 to either a random device or their counterpart. In this respect, we have made the following predictions:

Prediction 3: If an individual is given the opportunity to assign responsibility to another human or a random device, and choose to assign the responsibility to a random device, they will be less responsive in their rewards and punishments to their counterpart than if they were not given this opportunity.

Prediction 4: If an individual is given the opportunity to assign responsibility to another human or a random device, and choose to assign the responsibility to the human, they will be more responsive in their rewards and punishments to their counterpart than if they were not given this opportunity.

These predictions are based on Charness (1998), which found that assigning a wage by a random device dampened the responses from workers when compared to the wage being assigned by a third party. In view of this, we predict that players will react to the division in Stage 0 made by a random device similar to the manner in Charness' paper and relate the human decision more closely to how they responded to the third party decision-maker. Therefore, the random device in our experiment should dampen the responses.

Finally, we are interested in viewing how risk adversity affects the outcome of the task in Stage 0. Risk aversion is the viewed as an action made by an individual to avoid potential risk. In the "responsibility game" risk is introduced as the response given by a player's partner in the game as well as the unforeseen outcome of the random device if so chosen. There are several different aspects of the game that induce uncertainty to both the Players. Individuals will process each stage of the game to reflect their individual preferences with regards to risk. First, the Player (Player X) responsible for the decision in Stage 0, is faced with uncertainty of how the other Player (Y) will react to their decision, will Player Y respond negatively or positively to the choice. In addition, if Player X chooses to have a random device determine the allocations in Stage 1, there is uncertainty to the results of the random number draw.[23] This uncertainty, coupled with the impact of responsibility of the situation lead us to predict the following:

Prediction 4: Due to the impact of responsibility Player B will choose to have Player A make the decision over a random device.

Prediction 5: Due to the desire for a sense of control, Player A will choose to make the decision themselves more often then the random device.

C. PROCEDURE:

Participants were recruited for the experiment from Principles of Microeconomics and Macroeconomics courses offered at Virginia Tech University. The experiment itself took place in the Economics Laboratory at Virginia Tech. There the participants used a

[23] Note the expected payoff of the random device is 1.43 (or $0.715 USD).

computer interface to report their decisions in the "responsibility game". In total the experiment was conducted over 10 sessions totaling of 110 subjects.

In each session of the experiment, subjects entered the lobby of the experimental lab and were asked to fill out consent forms and wait for all expected subjects to arrive. They were then assigned randomly to computer terminals by selecting an index card from a stack that was labeled with a number corresponding to a specific computer terminal. The subjects were then asked to enter the lab and sit at their assigned computer terminal.

Prior to the experimental session, the computer terminals were assigned to either Player A or Player B and paper instructions were placed at each terminal respectively. Once the subjects were seated at their computer terminals, they were asked to silently read the instructions. After they had completed reading the directions, the subjects completed a brief quiz detailing the method in which their payments would be determined at the end of the game, due to the complexity of the payment method we felt this was a very important step to increase saliency (See Appendix C). Once all of the subjects had successfully completed the quiz, they logged onto their computers and began the "responsibility game."[24]

The decisions made by Player B were collected using the strategy method. The strategy method requires that Player B provide decisions in response to each possible offer made by Player A in Stage 1 of the game. Additionally, in Treatment A-Choose, Player B was asked to respond in both the case when Player A made the offer in Stage 1 as well as if a random device had made the offer. In this way, we were then able to observe the differences in behavior by each player dependent on the decision-making

[24] The "responsibility game" was conducted via a computer interface using "z-tree". We would like to thank Falk, Fehr, and Fischbacher for providing us with the shell of their constituent game to modify to fit the "responsibility game."

mechanism in Stage 1. Using the strategy method allowed us to maximize the data collected from each subject within the experiment, we were able to see both the change in response by Player B dependent on the offer in Stage 1 as well as the changes in response dependent on the decision-making mechanism in Stage 1. Furthermore, while completing the "responsibility game" tasks, participants were asked to report their feelings in response to each allocation of wealth offered by their counterpart. These answers assisted in the evaluation of an individual's perception of a 'fair' outcome in our data analysis. Using this method we were able to avoid making assumptions on why decisions were made the way they were with regards to fairness.

After completing the "responsibility game", subjects were told their payoffs from the game on the computer interface and were asked wait patiently for further instruction. Subjects then participated in a risk task. The risk task employed was the gamble task from in Eckel et al. (2008). In this task, subjects were asked to choose between two gambles (one riskier than the other). They were told that one of the gambles would be selected at the end of the experiment and a dice would be thrown to determine the winning outcome. They were also informed that they would be paid the amount if they won the gamble. As a final task, subjects were asked to complete a brief survey to collect information about the player's demographics as well as data on their trustworthiness, altruism levels, and perception of others.

Once the surveys were completed, subjects were called individually out to the lobby and paid the total amount from the risk-task and the "responsibility game." Points from the "responsibility game" were converted as follows: 2 points = $1 USD. At this point the subjects were free to leave the experiment.

Table IV: Player B's response to A's offer

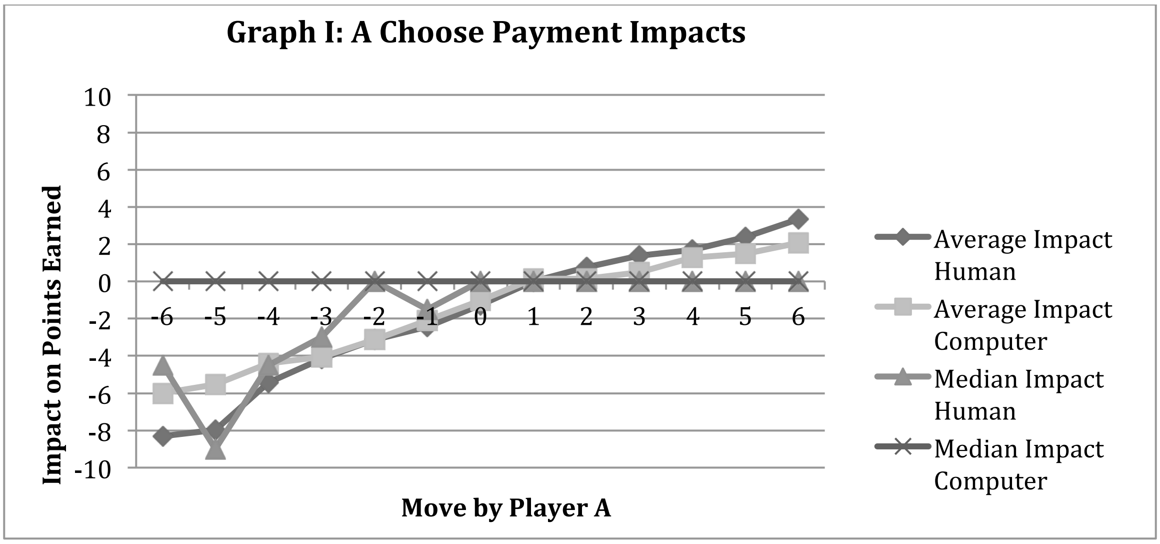
Treatment	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
A-Choose	-7.269	-6.058	-4.462	-4.115	-4.115	-2.769	-2.115	-1.038	0.019	0.481	1.096	1.596	2.577
	(8.303)	(6.941)	(5.751)	(5.016)	(4.278)	(3.924)	(3.705)	(3.923)	(4.104)	(4.548)	(4.811)	(5.340)	(6.140)
Human	-8.538	-6.923	-5.115	-4.308	-2.769	-2.308	-1.154	-0.731	0.500	1.000	1.192	2.000	3.077
	(8.472)	(7.088)	(5.935)	(4.978)	(4.236)	(4.097)	(3.802)	(4.153)	(4.042)	(4.534)	(5.154)	(5.727)	(6.517)
Computer	-6.000	-5.192	-3.808	-3.923	-2.769	-1.923	-0.923	-0.308	-0.462	-0.038	1.000	1.192	2.077
	(8.094)	(6.818)	(5.600)	(5.145)	(4.403)	(3.815)	(3.676)	(3.750)	(4.188)	(4.591)	(4.543)	(5.004)	(5.824)
B-Choose	-6.914	-6.362	-6.000	-5.000	-4.500	-3.621	-1.345	-0.328	-0.103	0.034	0.828	1.241	1.500
	(8.219)	(7.372)	(6.432)	(5.867)	(5.219)	(4.660)	(3.426)	(2.488)	(3.042)	(3.830)	(3.812)	(4.658)	(6.044)
Human	-7.379	-6.931	-6.310	-4.931	-4.448	-3.724	-0.931	-0.241	0.034	0.448	1.000	1.414	1.621
	(8.466)	(7.568)	(6.331)	(5.831)	(5.173)	(4.720)	(2.549)	(2.278)	(3.029)	(3.813)	(4.097)	(5.075)	(6.853)
Computer	-6.448	-5.793	-5.690	-5.069	-4.552	-3.517	-1.759	-0.414	-0.241	-0.379	0.655	1.069	1.379
	(8.087)	(7.257)	(6.628)	(6.006)	(5.356)	(4.680)	(4.129)	(2.719)	(3.101)	(3.868)	(3.568)	(4.284)	(5.233)

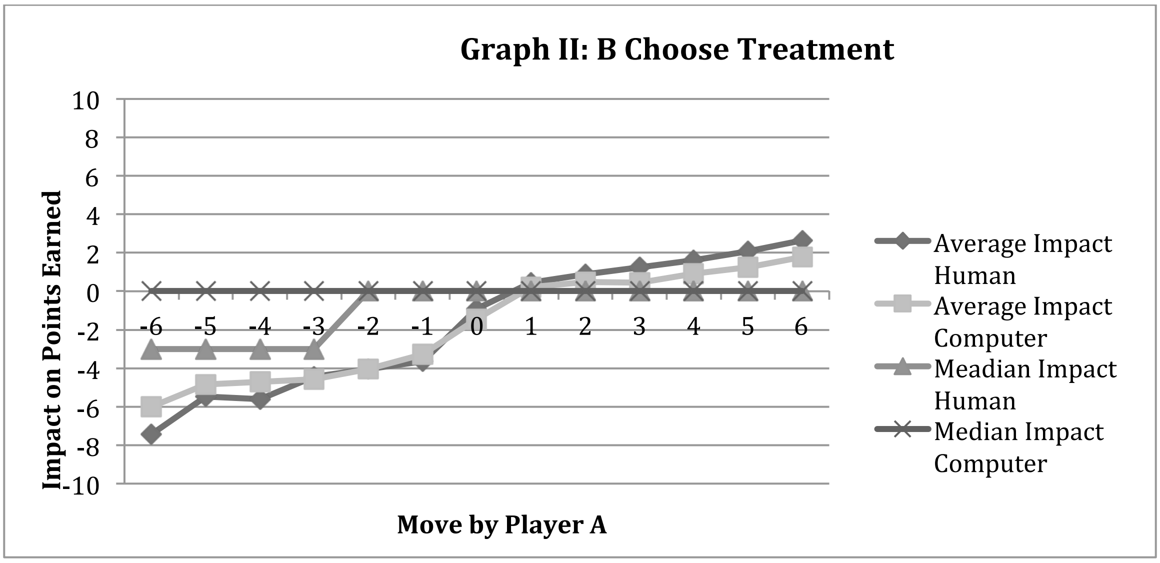
*Standard Deviation reported in Parentheses
*Human: reports the averages of the reponses by Player B to a decision made by a Human in Stage 1
*Computer: reports the averages of the reponses by Player B to a decision made by a Random device in Stage 1

4. RESULTS:

Table IV reports the summary statistics of moves made by Player B as a response to the offers made by Player A. Here we see that independent of treatment or intentionality, individuals punish offers made by Player A, which are negative and reward offers made by Player A that are negative. Upon further investigation we can see slight differences in the degrees to which Player B responds to positive and negative offers. The negative offers made by a Human Player A are punished to a slightly greater degree than those made by a Computer in both treatments. Furthermore, the positive offers made by a Human Player A are rewarded slightly higher than those made by a Computer Player A. Both of these results are independent of treatment, but as we see with further exploration, these results show that there is indeed a change in behavior caused by adding responsibility into the decision-making process.

The result that responsibility had an impact on the how the subjects behaved in the “responsibility game” is also apparent in Graphs I and II. In these Graphs Player B’s responses to Player A’s offers have been divided into these two graphs to visualize and isolate the effects of responsibility by treatment type as well as by decision-making mechanism. The graphs report the median and average impact on Player A from Player B dependent on Player A's decision. For example, as shown in Graph I, within the A choose treatment, a decision of 6 made by a human increased earnings a mean of 3.35 points and a median of 0 points while a decision of 6 made by the random device on behalf of Player A increased earnings a mean of 2.08 points and a median of 0 points. It is interesting to note that the median impact on Player A when a human made the decision in Stage 1 is non-zero in both the positive and negative realms. This is consistent throughout the B-Choose treatment as shown in Graph II and indicates that intentions within the "responsibility game" still influence the responses by Player B. However, the extent that Player B punishes and rewards Player A for their offers is smaller in the B-Choose treatment than in the A-Choose treatment.





Graphs I and II indicate that the extent to which Player B rewards Player A increases with the amount offered to Player B as one might expect and is seen in intentions literature. However, unlike Blount (1995) and Falk, Fehr, and Fischbacher (2008) we find that negative offers by Player A made on behalf of the random device tend to be punished relatively severely when compared to intentions literature, although not quite as strongly as those decisions made by Player A himself. Falk, Fehr, and Fischbacher's paper studies the effect of Player A's decision by assigning the decision-making task to either a random device or Player A. The experimenters made this decision prior to the any interactions between the players; therefore the results are studying the intentionality of offers made by Player A by removing the option to attribute responsibility. The also results indicate that in the A Choose treatment the fact that Player A chose to have the random device make the decision for them signals indirect responsibility which results in their punishment when offers are inequitable, even though that they do not have direct responsibility. When direct responsibility for the decisions in

Stage 1 is given to Player A, Player B responds more severely and punishes more for negative allocations.

Interestingly, the result that Player B punishes unfair random offers carries over into the B choose treatment where Player B's chose to have random offers made to them. However, Player B on average did not punish Player A as severely in the B Choose treatment as they did in the A Choose treatment. This is because Player B chose to have the random device and is therefore indirectly responsible for the outcome and act thusly in their responses. Player B depicted his realization of indirect responsibility by punished negative allocations made by a random device less severely than they do a human decision.

We also observe that positive offers made randomly in the B Choose treatment elicit rewarding behavior for Player A who made no decisions at all affecting their earnings. Player B confirms previous studies which show that it is fair that the random outcome is poor for them it is equitable that the outcome also be poor for player A and vice-versa if the outcome is favorable for Player B than it should also be favorable for Player A (Fehr 2006).

A. Impact on A's Payoff By Differences and Treatment

As indicated by Graphs I and II, attribution of responsibility does affect the outcome of the "responsibility game". To further explore the impact of attribution of responsibility we use an OLS regression model and employ differences of Player B's reaction as the dependent variable.

Table V
Regression with impact of B's Decision as a dependent variable differencing by Human - Computer

	I		II	III	IV
A Offer (a)		0.0904**	-0.130**	-0.133**	-0.134**
		(2.40)	(-2.35)	(-2.27)	(-2.28)
Treatment Dummy (c)		-0.13	-0.03	0.07	2.557*
		(-0.66)	(-0.10)	(0.21)	(1.88)
C x a			1.915***	1.959***	1.968***
			(9.68)	(8.65)	(8.59)
Gender				0.36	0.25
				(1.15)	(0.84)
Risk (r)				-0.15	0.07
				(-1.44)	(0.62)
Trust				0.04	0.04
				(1.04)	(1.04)
Reciprocity				0.00	0.00
				(-0.11)	(-0.12)
C x r					-0.361**
					(-2.08)
Constant	-0.0469	0.0146	0.01	0.72	(0.75)
	(-0.47)	(0.10)	(0.10)	(0.90)	(-0.83)
N	1430	1430	1430	1352	1352

t statistics in parentheses
* p<0.10, **p<0.05, *** p<0.01

One advantage of having Player B respond in a strategy method to both the offers made in Stage 1 as a human and by random device is that we can compare the differences in response levels on an individual level, as done in Table V. Table V reports four regressions with the impact of Player B's decision on Player A as the dependent variable differenced by decision mechanism. This variable differences the impact on Player A that Player B made in Stage 2 by subtracting the random device response counterpart from the human response. Regression I shows the impact of Player A's offer (a) on the differenced final payoff of Player A. We see that the offer made by Player A did have a significant impact on the decision made by Player B and thereby positively increases the final payoff of Player A if the decision was made by Player A. This finding is aligned with predictions 1 and 2, that independent of the decision-making mechanism, individuals will consistently sanction their counterpart for selfish behavior and reward their counterpart for kind behavior.

Next we note that the treatment does not have a significant impact on the response of Player B and therefore does not impact the payoff of Player A directly. Meaning that whether the decision in Stage 0 is made by Player A or Player B, there was no direct impact on the payoff of Player A. However, the interaction between the treatment dummy (T) and the offer (a) results in a statistically significant coefficient of 1.915 in regression III. Meaning that subjects were more responsive when offers were very high or very low in the A-Choose treatment, when the decisions were made by a human counterpart. This statistic further explains human behavior with regards to responsibility; individuals were more likely to reward behavior by Player A when the individual had taken direct responsibility for their actions and encourage this behavior through positive reinforcement.

By combining the conclusions of Falk, Fehr, and Fischbacher (2008) and these initial results, we see that the responsibility effect combines with the intentions effect and does significantly change the outcome of the game for Player A. To only address one aspect of decision-making would fail to appropriately model human behavior. Additionally, Falk, Fehr, and Fischbacher (2008) concluded that the role of intentions in the decision-making process does significantly effect the outcome with particular focus that the decisions made by a random device and negatively impacted Player B were not

Table VI:
Comparison regression with impact of B's decision as a dependent variable

	(I) Intentions	(II) Responsibility	(III) Intentions	(IV) Responsibility	(V) Intentions	(VI) Responsibility
A Offer (a)	0.295*	0.689***	0.278	0.658***	0.343	0.477***
	-1.88	-6.3	-1.2	-3.64	-1.44	-4.47
(H x a)	0.907***	0.181**	0.876**	0.350**	0.725**	0.0656
	-4.09	-2.4	-2.57	-2.51	-2.56	-0.7
Human Dummy (H)	-0.6157	-0.0937	-0.776	0.492*	0.149	0.258
	(-0.48)	(-0.47)	(-0.56)	-1.74	-0.15	-0.88
Constant	-0.401	-2.014***	-0.432	-2.287***	-0.57	-1.237***
	(-0.73)	(-4.05)	(-0.53)	(-3.51)	(-1.38)	(-2.77)
N	728	1430	336	660	392	770

t statistics in parentheses
* p<0.10, **p<0.05, *** p<0.01

° Columns I, III, and V are regressions using the data from Falk, Fehr and Fischbacher (2008); while Columns II, IV, and VI use data from Our integrity experiment. Columns I and II use all data in the positive and negative realm; Columns III and IV use only the positive offers from Player A and Columns V and VI use only data from the negative offers by Player A.

punished to the same degree as those made by a human. However, we see that once intentions are removed from the decision-making process (assigning the responsibility of decision-making to a random device) Player B consistently punishes their counterpart for negative impacts on their payoffs, because Player A is indirectly responsible.

In addition to the decision strategies made in the game, we collected information regarding the subjects risk aversion. Regression IV shows that there is no statistically significant impact of Player B's risk aversion (r) on the outcome of Player A by itself. However, we found that there is a statistically significant impact of the interaction between treatment type and risk aversion (T x r). If Player B is more risk averse in treatment A-choose, Player A's payoff is negatively impacted. In other words, if Player B attributes direct responsibility to Player A, Player B is less concerned about Player A's offers and punishes less and rewards more.

From the surveys completed by the subjects we found two distinct qualities that significantly impacted the payoff of Player A, trustworthiness and reciprocity.[25,26] As expected Player B's trustworthiness had a significant impact on Player A's payoff in the B Choose treatment, however was found to be insignificant in the A Choose treatment. Additionally, Player B's reciprocity score was found to have a significant impact on Player A's payoff in the A Choose treatment, but was irrelevant in the B Choose data. These results are intuitive and should be expected to show up in the data analysis, however psychological surveys are not commonly employed in experimental design. We believe that using these surveys can help seek out different aspects of human behavior that still need to be explored from an economic standpoint.

To better identify the impact of direct and indirect responsibility we ran regressions isolating the impact of responses by Player B on A by the negative and positive offers by Player A. We then compared our results to those of Falk, Fehr and Fischbacher (2008) to distinguish the impact of intentions from the impact of attribution of responsibility. These results are reported in Table VI where B's decision is used as the dependent variable. Regression II, IV, and VI are regressions using our data and regression I, III, and V report regressions using Falk, Fehr, and Fischbacher's (2008) data.

First, regressions I and II compare all of the data from our experiment directly with all of Falk, Fehr and Fischbacher's (2008) data. In comparison we see that when subjects were given the opportunity to attribute direct or indirect responsibility (in

[25] Trustworthiness and Reciprocity measures were calculated through surveys found in Robinson et al. (1991).

[26] Information about the subject's altruism was also collected. However, altruism and trustworthiness were found to be highly correlated (0.49) and we felt that analyzing the data using Trustworthiness was significantly more relevant.

regression II), the offer made by Player A has a larger impact on the responses by Player B compared to only attributions of intentions (in regression I). Meaning in both treatments, A Choose and B Choose, Player A's offer has an increasingly significant and greater impact on Player A's offer (in terms of response by Player B). However, in the intentions data from Falk, Fehr, and Fischbacher (2008) shown in regression II the interaction between Human Dummy (H) (Intentions Treatment) and offer has a more significant impact as well as a larger impact on Player A's payoff than in our data. This shows when only focusing on how intentions effect the outcome, individuals place a higher weight on the fact that a human making the decision in the higher and lower offers than when responsibility comes into play. For instance, if an individual was not able to select how a decision was made (ie the Falk, Fehr, and Fischbacher experiment), their counterpart (Player B) would be more responsive in Stage 2 to the actions in Stage 1, either by punishing more to a human or rewarding more to a human. On the other hand, given the opportunity to select how a decision was made (either by Player A or by Player B in our experiment) individuals are more concerned with the offer at hand over what decision-making mechanism was employed in Stage 1. We also see that by summing the impact from offer and interaction is equal to the total impact of the interaction from the Falk, Fehr, and Fischbacher (2008) results.

To further appreciate where these impacts are coming from we have separated the data out to look at the positive offers by Player A ($a \geq 0$) and the negative offers made by Player A ($a<0$). Columns III and IV report the analysis on the positive offers and

Columns V and VI report the analysis on the negative offers.[27] As shown above, Player B punishes in both the A Choose and B Choose datasets. However, this is done to a lesser extent when the offer made in Stage 1 is done by a random device rather than directly by Player A.

Focusing on the results from Player A's positive offers, in the "responsibility game" we still find that the offer has a higher and significant impact on Player B's response than in the Falk, Fehr, and Fischbacher paper. The response of Player B to Player A's positive offers in the Falk, Fehr, and Fischbacher paper are increasingly dependent on the interaction between Human offers and the degree of the offer. Meaning that individuals place a higher value on positive intentions when responsibility is removed from the scenario. Looking at columns V and VI, which report the negative offers by Player A, we see that the interaction between Human and offer are no longer significant in the "responsibility game." All of the impact from the interaction in the "responsibility game" comes from the positive offers, meaning that Player B does not place symmetric values over the positive and negative offers by Player A. The subjects in the "responsibility game" were more concerned with the negative offers than with the positive offers due to the attribution of responsibility albeit direct or indirect.

B. PATTERNS

By studying the behavior of Player B's response patterns we can better understand particular motivations of each player. To do this we have categorized Player B's behavior into four separate categories. First, we define to *Sanction* as the action pattern of taking

[27] Table VI reports the impact of the offers on Player A's payoff, noting this, we do no feel that it is appropriate to compare the impacts between the negative and positive offers because of the multiplier for the negative responses by Player B.

Table VII: B's Behavior Patterns by Category

	Sanction	Reward	No Pattern	Uniform	Same*
A Choose Treatment (n=52)					
All	38.46%	40.38%	23.08%	28.85%	69.23%
Computer	30.77%	34.62%	26.92%	30.77%	69.23%
Human	46.15%	46.15%	19.23%	26.92%	69.23%
Significance of Difference***	0.009	0.003	0.101	0.000	0.000
B Choose Treatment (n=58)					
All	39.66%	22.41%	18.97%	29.31%	72.41%
Computer	37.93%	13.79%	17.24%	34.48%	72.41%
Human	41.38%	31.03%	20.69%	24.14%	72.41%
Significance of Difference***	0.009	0.005	0.000	0.003	0.000
Both Treatments (n=110)					
All	39.09%	30.91%	20.90%	29.09%	70.91%
Computer	34.55%	23.64%	21.82%	32.73%	70.91%
Human	43.64%	38.18%	20.00%	25.45%	70.91%
Significance of Difference***	0.000	0.000	0.000	0.000	0.000

*We define Same as when Player B behaves in the same fashion with regards to any of the four categories for both the Human and the Computer (Note: This does not implay that the Players responses were of the same magnitude in each category)

**Description of Categories

*** Significane of difference between Human and Computer decision-making (Fischer's exact test, p-values)

away from Player A with respect to the offer presented. In other words, as the offers by Player A increase, the responses by Player B must decreases at some rate. Secondly, to *Reward* is defined by the action pattern of Player B gifting points to Player A in some pattern related to the amount offered by Player A. The third and fourth definitions relate to responses by Player B with no distinct pattern of increasing or decreasing with respect to offers made by Player A. If Player B's responses did not consistently increase or decrease with respect to offers made by A, we define this as *No Pattern*. Player B is categorized as *Uniform*, if Player B responds with a constant quantity for all offers made by Player A. An additional category which can engage in any of the four patterns above, is the *Same*, which is defined as when Player B did not differ in patterns when the decision made in Stage 1 were made by a Human or a random device.[28] The values in Table VI report that percentages that each of these patterns occurred in the experiment.

[28] The patterns of all the subjects are reported in Appendix A, Tables A and B.

The significance of difference reports the difference between Human and Random Device decision-making by the Fischer's exact test.

First we note that in all three categories, A Choose, B Choose and both treatments, the subjects were more likely to sanction and reward humans over computer respondents.[29] Here humans elicit a greater response in to either of the offer's extremes and reflect the reaction of Player B to increased responsibility of Player A, direct rather than indirect. Additionally, subjects were more likely to provide a uniform response to computer respondents over human counterparts.

There is a discrepancy between the rate of No Patterns between the A Choose and B Choose treatments. In the B Choose treatment we find that subjects were more likely to have no pattern with their human counterparts than with computer counterparts. However, in the A Choose treatment subjects were more likely to have no pattern with their computer counterparts than with their human counterparts.

A similar result to the 'No Intentions' treatment in Falk, Fehr, and Fischbacher (2008) was that 30% of the responses given to a computer by Player B resulted in no reciprocal behavior[30]. Therefore, in spite of the difference in attribution of responsibility, participants felt inclined to the same regularity to neither punish nor reward their counterpart if the allocation of wealth in the first stage was determined by a random device.

[29] It is important to remember that sanction and reward patterns here are not defined by extent to which the subjects rewarded and sanctioned, but only isolate the patterns that occurred.
[30] Player B did not sanction or rewards any offers made by Player A.

Table VIII
Regression with Impact of A's Decision as the dependent variable

Risk Aversion	1.296**
	(2.10)
A Choose Dummy (T)	2.91
	(1.30)
Trustworthiness	0.03
	(0.11)
Altruism	0.638*
	(1.98)
Male Dummy	-4.773**
	(-2.46)
Relative income	2.614**
	(2.45)
Constant	(5.05)
	(-1.00)
N	55

t statistics in parentheses
* p<0.10, **p<0.05, *** p<0.01

C. A'S IMPACT ON B'S PAYOFF

By collecting data about each individual participant in survey form we were able to run analysis on the decisions made by Player A in Stage 1. Table VII shows these results as a regression on impact of A's decision on B's initial standings in Stage 1 as the dependent variable. As anticipated, both risk and gender play a significant role on the offers made by Player A in Stage 1.[31] Individuals who are more risk averse were more inclined to give higher amounts to Player B in Stage 1. As shown in other literature, income also has a significant impact on the amount offered by Player A (reference). The regression shows that individuals with lower incomes are more inclined to give money to

[31] We did not find that there was a gender effect on risk aversion.

Player B. Player A did not react to being assigned the decision-making task in Stage 1 by being more kind in their offers to B. The offers made by Player A were consistent between treatment types.

D. ATTRIBUTION OF RESPONSIBILITY

In Stage 0 of the "responsibility game" either Player A or Player B was able to assign responsibility of the decision made in Stage 1 to a random device or to Player A. Within the survey, each subject was asked question to gather data about their trustworthiness in others, altruism, reciprocity and other socioeconomic data. We then used this the regression model below to determine how these variables impacted decisions made by Player A.

Assigning the decision to the random device represents to placing indirect responsibility to the decision-maker in Stage 0 and assigning the decision to Player A represents assigning direct responsibility to Player A. In the A-Choose treatment, 100% of the players chose to make the decision themselves, taking full and direct responsibility for their decisions. However, in the B-Choose treatment 50% of the participants chose to have the random device determine the division of wealth in Stage 1 and by definition took indirect responsibility upon them.

We interpret this in the following way, Player B viewed their decision in Stage 0 as both an assignment of responsibility and a signal of trust in Player A. When Player B did not attribute the responsibility to Player A, he was signaling distrust to Player B and felt that he would be better of choosing the random device. On the other hand, although the expected payoff of the random device was greater than 0, when Player B selected Player A to make the decision in Stage 1, they felt that the signal of trust could potentially increase their expected earnings.[32] Further emphasizing the effect responsibility has on choice. Conversely, when given the opportunity to choose, Player A always chose make the decision themselves and were thusly rewarded as shown above for kind allocations.

What we did not find however was a correlation between the expected feelings projected by Player A and their decision to make the offer themselves in the A Choose treatment. Differencing by Human feelings-Random Device, 21% of the time individuals thought their decisions made personally (by a Human) would result in a more positive feelings by their counterpart. Therefore, their decision to make the offer on their own accord is likely explained by their desire to hedge their risks and be in complete control of the outcome in Stage 1, rather than Player B's feelings toward the outcome.

E. Testing Theories of Fairness

Often models of fairness include variables that describe the feelings from both players about the potential outcome of a game. The feelings aspect of the interaction was determined by collecting survey responses about how each individual feels about the potential distributions of wealth. The survey has allowed us to test previously published theories' validity, such as Rabin's fairness model (Rabin M. , 1993).

Rabin (1993) modeled fairness based on the concept that individuals will behave in ways that they anticipate others to feel. In particular with respect to our game, Player A's offers would be directly impacted by how Player A expects Player B will feel about

[32] We found no difference in expected payoff between the offers made by a human and a random device in stage 0. Player A behaved consistently with those in Abbink et al (2000).

their potential offer. Table IX shows that predictions on other's feelings in fact does has a strong impact ($p=0.00$) on the offer made by Player A, agreeing with Rabin's model.

Table IX
Regression with A's offer as the dependent variable

Variable	Coefficient
A's Feelings	1.655***
	(11.30)
Constant	-0.46
	(-1.42)
N	55.00

t statistics in parentheses
* p<0.10, **p<0.05, *** p<0.01

5. CONCLUSIONS:

Our data consistently shows it is natural for humans to take into account both the acceptance of responsibility and the intentions behind individual behavior. Participants reflected their feelings about the other player's actions through rewards and punishments. In A-Choose treatment, participants playing the role of Player B punished both the random device and the human decisions when they were selfish. However, the punishments for the random device were not as strong as those for the human decision. We interpret this to show human's general admiration for taking responsibility for ones actions. If Player A was able to take responsibility for the outcome and they behaved selfishly they should be punished for their behavior, which is consistent with literature on attribution of intentions (Falk, Fehr, Fischbacher (2008), Blount (2000)). However, when Player A chooses to have a computer make the decision and selfish outcomes occur, Player B still punishes Player A. This result is contrary to intentions literature; in previously conducted studies decisions made by a random device were often not punished at all by Player B for the outcome (Falk, Fehr, and Fischbacher (2008)). Again, this shows that there is a significant reaction to the context of responsibility that needs to be addressed in utility theory. Individuals value responsibility in a decision making task and if responsibility is forgone and bad things happen, Player A was still held accountable for the outcomes because they had indirect responsibility for choosing the device that allocated the endowments.

While our experiment introduces responsibility into a social interaction, it focuses on anonymous exchanges between individuals. Often employment of a random decision making device is used to remove responsibility from an interaction because tense interactions. For instance, Alexander Sebald (2009) describes a scenario between three friends; friends A and B are arguing over an extra concert ticket that friend C has to allocate to one of them. Rather than making the decision himself, friend C elects to flip a coin, removing the responsibility of the decision-making task and removing himself from inevitable backlash from the ticketless friend. Our experiment differs from this scenario in several ways, but most importantly it is by the familiarity and repeated interaction amongst friends that result a difference in reactions. Our experiment addresses anonymous interactions and decision-making in a one shot game. The question that we are looking to answer: Does taking responsibility for one's actions change the response within an anonymous interaction? As well as, how does the attribution of responsibility affect the response within an anonymous interaction? In effect though our experimental game we are examining the potential effects of a responsibility signal (the choice between a randomly generated offer and a human offer) on the outcome of the game.

7. References

Abbink, K, and B Irlenbusch…. "The moonlighting game:: An experimental study on reciprocity and retribution." *Journal of Economic Behavior & …*, Jan 2000.

Barnes, W, and W Falk…. "Symposium: Intention, Motive and Responsibility." *Proceedings of the Aristotelian …*, Jan 1945.

Bentham,. 1970. "An Introduction to the Principals of Morals and Legislation"

Blount, S. "When Social Outcomes Aren′ t Fair: The Effect of Causal Attributions on Preferences." *Organizational behavior and human decision …*, Jan 1995.

Bonetti, S. "Experimental economics and deception." *Journal of Economic Psychology*, Jan 1998.

Charness…, G. "Responsibility-Allevation and Effort Provision in a Gift-Exchange Experiment." *econ.upf.edu*, Jan 1998.

Charness, G., 2000. Responsibility and effort in an experimental labor market. Journal of Economic Behavior and Organization 42, 375–384.

Charness, Gary, and Matthew O. Jackson**.** 2009. "The Role of Responsibility in Strategic Risk-Taking." *Journal of Economic Behavior and Organization*.

Charness, Gary, Masclet, David and Villeval, Marie-Claire, Competitive Preferences and Status as an Incentive: Experimental Evidence. IZA Discussion Paper No. 5034

Dufwenberg, M., Kirchsteiger, G., 2004. "A theory of Sequential Reciprocity." *Games and Economic Behavior*. 47 (2), 268-298

Eckel…, C. "Forecasting risk attitudes: An experimental study using actual and forecast gamble choices." *Journal of Economic Behavior & Organization*, Jan 2008.

Falk, A, and E Fehr…. "Testing theories of fairness--Intentions matter." *Games and Economic Behavior*, Jan 2008.

Falk, A., Fischbacher, U., 2006. "A theory of reciprocity." Games Econ. Behav. 54 (2), 293–315.

Heider, F. 1958. "The Psychology of Interpersonal Relations." *John Wiley and Sons, New York.*

Hoffman, E. and Spitzer, M. 1985 "Entitlements, Rights, and Fairness: An Experimental Examination of Subjects' concepts of Distributional Jusitice." *Journal of Legal Studies*, vol XIV

Holt…, C. "Risk aversion and incentive effects." *American Economic Review*, Jan 2002.
Pound, R. "The Economic Interpretation and the Law of Torts." *Harvard Law Review*, Jan 1940.

Kelly. H.H. 1967 "Attribution in Social Psychology." *Nebraska Symposium on Motivation* 15, 192-238.

Kelly, H.H., 1973 "The Processes of Causual Attribution." *American Psychologist* 28: 107-128.

Rabin, M. "Risk aversion and expected-utility theory: A calibration theorem." *Econometrica*, Jan 2000.

Rabin, Matthew. "Incorporating Fairness into Game Theory and Economics." *The American Economic Review* 83, no. 5 (Dec 1993): 1281-1302.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. 1991. Measures of personality and social psychological attitudes. San Diego: Academic Press.

Ross, M., Fletcher, G.J.O., 1985. "Attribution and Social Perception." *Lindsey, G. Aronson, E. (Eds.) The Handbook of Social Psychology*, Vol. 2. Random House, NY, pp73-114

Sebald, A. "Attribution and reciprocity." *Games and Economic Behavior*, Jan 2010.
Shultz, T, and K Wright…. "Assignment of moral responsibility and punishment." *Child Development*, Jan 1986.

8. Appendix

A: Sanction and Reward

A. Sanctions and Reward Patters by Player B in Treatment A Choose																		
Mechanism	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**Sanction**	**Reward**	**No Pattern**	**Uniform**	**Same**
Computer	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	-18	-18	-18	-18	-18	-18	-18	-18	-18	-18	-18	-18	-18				x	x
Human	-18	-18	-18	-18	-18	-18	-18	-18	-18	-18	-18	-18	-18				x	x
Computer	-18	-15	-9	-9	0	0	0	0	0	0	0	0	0	x	x			x
Human	-18	-15	-9	-6	0	0	0	0	2	3	4	7	10	x	x			x
Computer	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	0	-3	0	-3	0	-3	0	1	2	3	4	5	8			x		x
Human	0	-3	0	-3	0	-3	0	1	2	3	4	5	8			x		x
Computer	-12	-9	-9	-9	-6	-3	0	1	1	3	5	7	9	x	x			x
Human	-18	-15	-12	-9	-6	-3	0	2	4	6	8	10	12	x	x			x
Computer	-15	-12	-9	-9	-6	-6	0	2	3	4	4	4	5	x	x			x
Human	-15	-12	-9	-6	-3	0	0	0	1	2	3	3	4	x	x			x
Computer	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	-18	-18	-12	-9	-6	-3	0	2	4	6	8	10	14	x	x			x
Human	-18	-18	-12	-9	-3	-3	0	2	4	6	8	10	14	x	x			x
Computer	-6	-12	0	-9	-3	1	0	2	-6	-9	4	2	0			x		x
Human	-9	-6	2	2	-6	-12	0	-3	2	4	-6	-6	0			x		x
Computer	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	0	0	0	0	0	0	0	0	0	1	1	1	1		x			x
Human	0	0	0	0	0	0	0	0	1	1	1	1	1		x			x
Computer	0	0	-3	-3	0	0	0	-3	-6	-3	0	-3	0			x		x
Human	0	-3	-3	-6	0	0	-6	-3	0	-3	-3	0	-3			x		x
Computer	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	-18	-15	-12	-9	-6	-3	0	1	2	4	6	8	10	x	x			x
Human	-18	-15	-12	-9	-6	-3	0	1	2	4	6	8	10	x	x			x
Computer	-18	-3	-3	-3	-3	-3	0	1	1	1	1	1	6	x	x			x
Human	-18	-3	-3	-3	-3	-3	0	1	1	1	1	1	1	x	x			x
Computer	0	0	0	0	0	0	0	1	2	3	4	5	6		x			x
Human	0	0	0	0	0	0	0	1	2	3	4	5	6		x			x
Computer	0	0	0	0	0	0	0	0	0	0	1	0	0			x		
Human	-18	-15	-6	-3	0	0	0	0	0	0	0	0	1	x	x			
Computer	0	0	0	0	0	0	0	0	0	0	1	3	4		x			
Human	0	0	0	0	0	-3	0	1	0	0	2	4	6			x		
Computer	0	0	0	0	-3	-3	0	2	3	4	5	6	9			x		
Human	-15	-12	-12	-9	-6	-3	0	2	4	5	5	7	10	x	x			
Computer	-15	-12	-12	-12	-12	-6	-6	0	0	0	0	0	0	x				
Human	-15	-12	-12	-12	-9	-3	-6	-9	0	0	0	0	0			x		
Computer	0	0	0	0	0	0	0	0	0	0	0	0	0				x	
Human	-18	-15	-12	-9	-6	-3	0	2	4	6	8	10	12	x	x			
Computer	-18	-15	-12	-9	-6	-3	0	0	0	0	0	0	0	x				
Human	-18	-15	-12	-9	-6	-3	0	1	2	3	4	5	6	x	x			
Computer	0	0	0	0	0	0	0	0	0	0	0	0	0			x		
Human	-3	-3	-3	-3	0	0	0	0	0	0	0	0	0	x				
Computer	0	-3	0	0	-3	0	0	0	0	0	0	0	0			x		
Human	-3	0	0	0	0	0	0	0	0	0	0	0	0	x				

B. Sanctions and Reward Patters by Player B in Treatment B Choose

Mechanism	Choice	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	Sanction	Reward	No Pattern	Uniform	Same
Computer	x	-6	-6	-6	-3	-6	-3	-6	-3	-3	-9	0	0	0			x		x
Human		-12	-6	-6	-3	-6	-9	-3	-3	0	-6	0	0	0			x		x
Computer		-18	-18	-15	-15	-15	-12	0	0	0	0	0	0	0	x				x
Human	x	-18	-18	-15	-15	-15	-12	-12	0	0	0	0	0	0	x				x
Computer	x	-12	-9	-9	-9	-9	-9	0	0	0	0	0	0	0	x				x
Human		-9	-9	-9	-9	-9	-9	0	0	0	0	0	0	0	x				x
Computer	x	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human		0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer		-18	-15	-12	-9	-6	-3	0	0	0	0	0	0	0	x				x
Human	x	-18	-15	-12	-9	-6	-3	0	0	0	0	0	0	0	x				x
Computer		2	3	-6	-15	-12	-9	-15	4	5	1	4	7	10			x		x
Human	x	2	3	-12	-9	-12	-9	-6	3	5	7	4	7	10			x		x
Computer		-18	-18	-18	-15	-15	-15	0	0	0	0	0	0	0	x				x
Human	x	-18	-18	-18	-15	-15	-15	0	0	0	0	0	0	0	x				x
Computer	x	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human		0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	x	-6	-6	-6	-6	-6	-6	0	0	0	0	0	0	0	x				x
Human		-6	-6	-6	-6	-6	-6	0	0	0	0	0	0	0	x				x
Computer		-18	-15	-12	-9	-6	-3	0	2	4	6	8	10	12	x	x			x
Human	x	-18	-15	-12	-9	-6	-3	0	2	4	6	8	10	12	x	x			x
Computer	x	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human		0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	x	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human		0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer		0	0	0	0	0	0	0	-3	-3	-3	-6	-6	-9	x				x
Human	x	0	0	0	0	0	0	0	-3	-3	-3	-3	-3	-15	x				x
Computer	x	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human		0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	x	0	0	0	0	-9	-6	-3	0	0	0	0	0	0			x		x
Human		0	-15	-12	-9	-9	-3	-3	0	0	0	0	0	0			x		x
Computer	x	-18	-15	-18	-6	-6	-3	0	0	-3	-6	0	0	0			x		x
Human		-18	-18	-6	1	-3	-3	0	-3	-6	0	0	0	0			x		x
Computer	x	0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Human		0	0	0	0	0	0	0	0	0	0	0	0	0				x	x
Computer	x	0	0	0	0	0	0	0	0	0	1	2	3	4		x			x
Human		0	0	0	0	0	0	0	0	0	1	2	3	4		x			x
Computer		-18	-15	-12	-9	-6	-3	0	0	0	0	0	0	0	x				x
Human	x	-18	-15	-12	-9	-6	-3	0	0	0	0	0	0	0	x				x
Computer		-9	-12	-15	-15	-12	-6	-3	2	3	5	7	9	11			x		x
Human	x	-9	-12	-15	-15	-9	-6	-3	2	3	5	7	9	11			x		x
Computer	x	0	0	0	0	0	0	0	2	4	6	8	10	12		x			x
Human		0	0	0	0	0	0	0	2	4	6	8	10	12		x			x
Computer		0	0	0	0	0	0	0	0	0	0	0	0	0				x	
Human	x	0	0	0	0	0	0	0	0	0	1	1	2	2		x			
Computer	x	-18	-15	-12	-12	-12	-12	-12	-9	-9	-9	-6	-6	-6	x				
Human		-18	-15	-12	-12	-12	-12	0	2	3	3	4	4	5	x	x			
Computer	x	0	0	0	0	0	0	0	0	0	0	0	0	0				x	
Human		0	0	0	0	0	0	0	1	2	3	4	5	6		x			
Computer		0	0	0	0	0	0	0	0	0	0	0	0	0				x	
Human	x	-18	-15	-12	0	0	0	0	0	0	0	0	0	6	x	x			
Computer		0	0	0	0	0	0	0	2	4	6	8	10	12		x			
Human	x	0	0	0	0	0	-3	0	2	4	6	8	10	12	x	x			
Computer		-12	-12	-9	-9	0	0	0	0	0	0	0	0	0	x				
Human	x	-18	-12	-9	-9	-3	0	0	0	0	2	4	5	6	x	x			
Computer	x	-18	-15	-15	-15	-12	-12	-12	-9	-9	-9	-6	-6	-6	x				
Human		-18	-15	-15	-15	-12	-12	0	-9	-9	-9	-6	-6	-6				x	
Computer	x	0	0	0	0	0	0	0	0	0	0	0	0	0				x	
Human		0	0	0	0	0	0	0	-3	-6	-9	-12	-15	-18			x		

B: GRAPHS

Table I.a
Regression with impact of B's decision as a dependent variable

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
A Offer (a)	0.779***	**0.689*****	0.779***	0.689***	**0.689*****	**0.664*****	0.664***
	-6.94	**-6.3**	-6.94	-6.29	**-6.29**	**-6.05**	-6.05
Human Dummy (H)	-0.0937	**-0.0937**	-	-0.0937	**-0.0236**	**1.505**	1.48
	(-0.47)	**(-0.47)**		(-0.47)	**(-0.08)**	**-1.08**	-1.08
(H x a)		**0.181****	-	0.181**	**-0.158**	**-0.170***	-0.165
		-2.4		-2.4	**(-1.62)**	**(-1.69)**	(-1.65)
A Choose Dummy (T)			0.132	0.132	**0.206**	**0.267**	-5.003
			-0.14	-0.14	**-0.25**	**-0.31**	(-1.10)
(T x a)					**1.474*****	**1.519*****	1.498***
					-6.6	**-6.8**	-6.79
B's Gender (Male==1)						**-0.414**	-0.172
						(-0.53)	(-0.21)
B's Risk Aversion (r)						**0.234**	-0.242
						-0.78	(-0.40)
B's Trustworthiness						**-0.0339****	-0.0340**
						(-2.06)	(-2.07)
B's Reciprocity						**-0.0311***	-0.0310*
						(-1.96)	(-1.96)
(r x T)							0.766
							-1.18
(r x H)						**-0.208**	-0.204
						(-1.16)	(-1.16)
Constant	-2.014	**-2.014*****	-2.123***	-2.076***	**-2.112*****	**-3.837***	-0.711
	(-4.06)	**(-4.05)**	(-3.49)	(-3.28)	**(-3.29)**	**(-1.89)**	(-0.18)
N	1430	**1430**	1430	1430	**1430**	**1352**	1352

t statistics in parentheses
* p<0.10, **p<0.05, *** p<0.01

Table I.b

A Choose Treatment: Regression with impact of B's decision as a dependent variable

	(I)	(II)	**(III)**	**(IV)**	**(V)**	**(VI)**
A Offer (a)	0.774***	0.640***	0.636***	**0.636*****	0.636***	**0.637*****
	-4.62	-3.9	-3.72	**-3.72**	-3.72	**-3.74**
Human Dummy (H)	-0.231	-0.231	-0.0954	**-0.0954**	0.461	**0.37**
	(-0.91)	(-0.91)	(-0.43)	**(-0.43)**	-0.5	**-0.44**
(H x a)		0.269**	0.225*	**0.225***	0.225*	**0.225***
		-2.18	-1.88	**-1.87**	-1.87	**-1.86**
B's Gender (Male==1)			-1.09	**-0.776**	-0.776	**-0.776**
			(-0.86)	**(-0.66)**	(-0.66)	**(-0.66)**
B's Risk Aversion (r)				**0.588****	0.628**	**0.622****
				-2.17	-2.31	**-2.3**
(r x H)					-0.0808	**-0.068**
					(-0.63)	**(-0.57)**
B's Trustworthiness						**-0.0548****
						(-2.24)
B's Reciprocity						**-0.0056**
						(-0.18)
Constant	-1.876**	-1.876**	-1.324**	**-5.568****	-5.846**	**-5.756****
	(-2.47)	(-2.47)	(-2.19)	**(-2.61)**	(-2.72)	**(-2.75)**
N	676	676	650	**650**	650	**650**

t statistics in parentheses

* p<0.10, **p<0.05, *** p<0.01

Table I.c

B Choose Treatment: Regression with impact of B's decision as a dependent variable

	(I)	(II)	(III)	(IV)	(V)	(VI)
A Offer (a)	0.784***	**0.733*** **	0.690***	**0.690*** **	0.690***	**0.689*** **
	-5.09	**-4.92**	-4.81	**-4.81**	-4.8	**-4.78**
Human Dummy (H)	0.0292	**0.0292**	0.0228	**0.0228**	-1.389	**-1.468**
	-0.1	**-0.1**	-0.07	**-0.07**	(-0.72)	**(-0.80)**
(H x a)		**0.102**	0.11	**0.11**	0.11	**0.108**
		-1.12	-1.13	**-1.13**	-1.13	**-1.11**
B's Gender (Male==1)			0.18	**0.461**	0.461	**0.461**
			-0.15	**-0.35**	-0.35	**-0.35**
B's Risk Aversion (r)				**-0.418**	-0.521	**-0.529**
				(-0.66)	(-0.81)	**(-0.83)**
(r x H)					0.206	**0.221**
					-0.78	**-0.88**
B's Trustworthiness						**-0.0326**
						(-1.46)
B's Reciprocity						**-0.0481** **
						(-2.30)
Constant	-2.138***	**-2.138*** **	-2.326***	**0.375**	1.081	**0.736**
	(-3.22)	**(-3.22)**	(-3.32)	**-0.09**	-0.26	**-0.18**
N	754	**754**	702	**702**	702	**702**

t statistics in parentheses

* p<0.10, **p<0.05, *** p<0.01

C: INSTRUCTIONS FOR PLAYER A CHOOSE TREATMENT

C.I:

……………………………………………………………………………………………

Instructions for Participants A

You are now taking part in an economic experiment. Please read the following instructions carefully. You will learn everything you need to know in order to participate in this experiment. Don't hesitate to ask, if there is anything you don't understand. Questions will be answered at your seat.

During the experiment you can earn **points**. The number of the points, which you will actually receive, depends on both your decisions and the decisions of the other participants.

At the end of the experiment all points earned during the experiment are converted into dollars at a rate of

2 Points = 1 Dollar

At the end of the experiment you will be paid the income you earned during the experiment.

Please note that during the whole experiment communication is not allowed. We also emphasize that you may activate only those functions at the computer that are related to the experiment. Any communication or 'playing around' with the computer will lead to the exclusion from the experiment. If you have questions, do not hesitate to ask us.

Description of the Experiment

1. At the beginning of the experiment, all 12 participants are randomly divided into 6 participants of type A and 6 participants of type B. **During the whole experiment you will be a participant of type A.**

2. Subsequently, the computer randomly arranges **groups of two** consisting of a participant A and a participant B. No participant knows with whom he is in a group, i.e., all decisions are made anonymously.

3. Each participant A and each participant B receives an **endowment of 12 points**.

4. The task consists of **three stages**:

   Stage 0: Participant A will decide whether he or she wants to make his or her own decision in stage 2 of the experiment, or let a computer choose for him or her.

   Stage 1: Either Participant A or the computer will make a move on behalf of Participant A.

   Stage 2: Participant B will make his or her decision.

These stages will now be described in more detail.

Stage 0:

In this stage you choose whether you to make your own decision in stage 1 of the experiment, or let a computer choose for your. The stage 1 task consists of picking a number between -6 and 6. If you choose to have the computer make your decision it generates a random number between 0-99 and then makes a decision according to the following table:

Realized number	*A*'s move *a*	Percent
0–6	−6	7
7–8	−5	2
9–15	−4	7
16–19	−3	4
20–21	−2	2
22–26	−1	5
27–39	0	13
40–46	1	7
47–55	2	9
56–62	3	7
63–73	4	11
74–75	5	2
76–99	6	24

Stage 1:

At this stage you will either make a decision or let the computer make it for you. The following explanation is written as if you have decided to make your own decision, but the payoff calculations are the same in either case. You can

a) either **increase** participant B's points,
b) or **decrease** participant B's points,
c) or leave participant B's points **unchanged**.

Case a)
You may **increase** participant B's points. In doing so you will have to bear **costs**. If you spend **one** point, participant B's points increases by **three** points. Suppose for example, that you spend **three** points - then your points decreases by **three** points, whereas participant B's increases by **9** points.

Case b)
You may also **decrease** participant B's points in order to increase your own points. Each point that you transfer from your points to participant B decreases participant B's points by **one** point and increases your points by **one** point. If, for instance, you decrease participant B's points by **3** points, your points increases by **3** points.

Case c)
Finally, you may leave your points **unchanged**. In this case participant B's points will remain unchanged, too.

All three possibilities are shown in the table below: The column "your decision" shows all **possible decisions** for you. Please note that an **increase** of participant B's points is indicated by **positive numbers**, whereas a **decrease** of participant B's points is indicated by **negative numbers**. The maximum by which participant B's points can be decreased is **6** points. This means that you may choose a number between **-6** and **-1**. If you wish to increase participant B's points, you choose a number between **+1** and **+6**. If you wish to leave participant B's points unchanged, you choose **0**.

The last two columns indicate the points at the end of stage 1, which result from different decisions.

	Your decision	Your points at the end of stage 1	Participant B's points at the end of stage 1
You increase B's points which decreases your points (Case a)	**+6**	12 - 6 = **6**	12 + 18 = **30**
	+5	12 - 5 = **7**	12 + 15 = **27**
	+4	12 - 4 = **8**	12 + 12 = **24**
	+3	12 - 3 = **9**	12 + 9 = **21**
	+2	12 - 2 = **10**	12 + 6 = **18**
	+1	12 - 1 = **11**	12 + 3 = **15**
(Case c)	**0**	12 + 0 = **12**	12 + 0 = **12**
You decrease B's Points which increases your points (Case b)	**-1**	12 + 1 = **13**	12 - 1 = **11**
	-2	12 + 2 = **14**	12 - 2 = **10**
	-3	12 + 3 = **15**	12 - 3 = **9**
	-4	12 + 4 = **16**	12 - 4 = **8**
	-5	12 + 5 = **17**	12 - 5 = **7**
	-6	12 + 6 = **18**	12 - 6 = **6**

Stage 2:

After you have made your decision at stage 1, both the points of A and B have **changed**. (The only instance where this does not hold is if you leave both points unchanged, i.e., in case c). At the second stage it is participant B's turn to make decisions. Starting from the changed points participant B can decide to

a) either **increase** your points that exists at the end of stage 1,
b) or **decrease** your points that exists at the end of stage 1,
b) or leave both points, as they exist at the end of stage 1, **unchanged**.

Case a)
Participant B may **increase** your points that exists at the end of stage 1. This is **costly** for participant B. Each point by which your points is increased will reduce participant B's points by one **point**. Suppose, e.g., participant B increase your points by 3 points; then participant B's own points is also reduced by 3 points.

Case b)
Participant B may also **reduce** your points that exists at the end of stage 1. This is also **costly** for participant B. Each point participant B spends reduces your points by **three** points. Thus, if participant B spends 3 points, her own points decreases by 3 points and your points decreases by 9 points.

Case c)
Participant B may also leave your points available at the end of stage 1 **unchanged**. In this case participant B's points from stage 1 remains unchanged, too.

The following table summarizes all **possible** decisions participant B can make. The column "participant B's decision" shows all possible decisions. An **increase** of your points is achieved by the choice of a **positive** number, a **decrease** by the choice of a **negative** number. Please note that an increase of your points is possible up to a maximum of 18 points. This means that participant B may choose a number between **+1** and **+18**. If participant B wishes to decrease your points, she can choose a number between **-1** and **-6**. If participant B wishes to leave the points unchanged, she can choose **0**.

As soon as participant B's decisions are made at stage 2, the **final** points are determined. The final points result from the points at the end of stage 1 and from participant B's decisions. The table shows these final points for you and participant B. The symbol "X" represents the points at the end of stage 1.

	Participant B's decision	Your final points (X = points at the end of stage 1)	Participant B's final points (X = points at the end of stage 1)
Participant B increases your points available at the end of stage 1 (Case a)	+18	X + 18	X - 18
	+17	X + 17	X - 17
	+16	X + 16	X - 16
	+15	X + 15	X -15
	+14	X + 14	X - 14
	+13	X + 13	X - 13
	+12	X + 12	X - 12
	+11	X + 11	X - 11
	+10	X + 10	X - 10
	+9	X + 9	X - 9
	+8	X + 8	X - 8
	+7	X + 7	X - 7
	+6	X + 6	X - 6
	+5	X + 5	X - 5
	+4	X + 4	X - 4
	+3	X + 3	X - 3
	+2	X + 2	X - 2
	+1	X + 1	X - 1
Case c)	0	X + 0	X - 0
Participant B decreases your points available at the end of stage 1 (Case b)	-1	X - 3	X - 1
	-2	X - 6	X - 2
	-3	X - 9	X - 3
	-4	X - 12	X - 4
	-5	X - 15	X - 5
	-6	X - 18	X - 6

How do you make your decision?

If you choose to have your decision made randomly you will not make a decision in stage 1. Otherwise decisions are made on the computer. The screen where you insert your decision at stage 1 looks as follows:

Your decision as a participant A	Your points at the end of the first stage	Participant B's points at the end of the first stage
☐ -6	18	6
☐ -5	17	7
☐ -4	16	8
☐ -3	15	9
☐ -2	14	10
☐ -1	13	11
☐ 0	12	12
☐ +1	11	15
☐ +2	10	18
☐ +3	9	21
☐ +4	8	24
☐ +5	7	27
☐ +6	6	30

OK

In the first column "Your Decision as Participant A" you see all possible decision between **-6** and **+6** you can make. In columns two and three ("Your points at the end of the second stage" and "Participant B's points at the end of the second stage") you see which points at the end of the second stage result from these possible decisions for you and Participant B. Please indicate which decision you make by checking the box next to the number of points you wish to transfer to or from Participant B. Remember that a positive number increases Participant B's points by the three times as much as it decreases your points. A negative number decreases Participant B's points by the same amount as it increase your points. When you have made your decision please press "OK"

Only after you have made your decision, the computer informs you which decision "your" Participant B has made. The final points for you and "your" Participant B is determined by the decision of participant B and the number you have inserted in the corresponding row. If, for example, if you choose +3 and participant B chooses +4, the final split is

for you: 12 – 3 + 4 = 13
for Participant B: 12 + (3 * 3) – 4 = 17

At the end of the experiment all participants will learn their counterpart's decision and their total earnings in the experiment.

The experimenters have previously carried out the experiment in which you are now participating. Participant A's in this former experiment have made certain decisions. The following table shows the percentage of the respective decisions made by participants A **in a previous experiment**. Because we have changed the experiment somewhat the results of this experiment are likely to be different.

Decisions of Participant A's in a previous experiment:

Decision	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	**+4**	**+5**	**+6**
Percentage of Participant A's who have chosen the corresponding decision	7 %	2 %	7 %	4 %	2 %	5 %	13 %	7 %	9 %	7 %	11 %	2 %	24 %

Do you have any questions? If you do please raise your hand and someone will be around to answer them. If not, please go ahead and answer the questions in the short quiz on the next two pages.

Quiz

Please calculate the following exercises before the experiment starts. Wrong answers have no consequences for you. The experiment will begin as soon as all participants have correctly solved the problems. When you are done, please consider which decisions you wish to make during the experiment.

1. Stage 1: participant A chooses **+5**.
   Stage 2: participant B chooses **+18**.

   What is participant A's point balance at the end of stage 1? ....
   What is participant B's point balance at the end of stage 1? ....

   What is participant A's final number of points? ....
   What is participant B's final number of points? ....

2. Stage 1: participant A chooses **0**.
   Stage 2: participant B chooses **0**.

   What is participant A's point balance at the end of stage 1? ....
   What is participant B's point balance at the end of stage 1? ....

   What is participant A's final number of points? ....
   What is participant B's final number of points? ....

3. Stage 1: participant A chooses **-6**.
   Stage 2: participant B chooses **-6**.

   What is participant A's point balance at the end of stage 1? ....
   What is participant B's point balance at the end of stage 1? ....

   What is participant A's final number of points? ....
   What is participant B's final number of points? ....

4. Stage 1: participant A chooses **-3**.
   Stage 2: participant B chooses **-1**.

   What is participant A's point balance at the end of stage 1? ....
   What is participant B's point balance at the end of stage 1? ....

   What is participant A's final number of points? ....
   What is participant B's final number of points? ....

5. Stage 1: participant A chooses **-3**.
Stage 2: participant B chooses **+1**.

What is participant A's point balance at the end of stage 1? ....
What is participant B's point balance at the end of stage 1? ....

What is participant A's final number of points? ....
What is participant B's final number of points? ....

6. Stage 1: participant A chooses **+4**.
Stage 2: participant B chooses **-2**.

What is participant A's point balance at the end of stage 1? ....
What is participant B's point balance at the end of stage 1? ....

What is participant A's final number of points? ....
What is participant B's final number of points? ....

C.II

....

Instructions for Participants B

You are now taking part in an economic experiment. Please read the following instructions carefully. You will learn everything you need to know in order to participate in this experiment. Don't hesitate to ask, if there is anything you don't understand. Questions will be answered at your seat.

During the experiment you can earn **points**. The number of the points, which you will actually receive, depends on both your decisions and the decisions of the other participants.

At the end of the experiment all points earned during the experiment are converted into dollars at a rate of

2 Points = 1 Dollar

At the end of the experiment you will be paid the income you earned during the experiment.

Please note that during the whole experiment communication is not allowed. We also emphasize that you may activate only those functions at the computer that are related to the experiment. Any communication or 'playing around' with the computer will lead to the exclusion from the experiment. If you have questions, do not hesitate to ask us.

Description of the Experiment

1. At the beginning of the experiment, all 12 participants are randomly divided into 6 participants of type A and 6 participants of type B. **During the whole experiment you will be a participant of type B.**

2. Subsequently, the computer randomly arranges **groups of two** consisting of a participant A and a participant B. No participant knows with whom he is in a group, i.e., all decisions are made anonymously.

3. Each participant A and each participant B receives an **endowment of 12 points**.

4. The task consists of **three stages**:

   Stage 0: Participant A will decide whether he or she wants to make his or her own decision in stage 1 of the experiment, or let a computer choose for him or her.

   Stage 1: Either Participant A or the computer will make a move on behalf of Participant A.

   Stage 2: You will make your decision.

These stages will now be described in more detail.

Stage 0:

In this stage participant A will choose whether to make his or her own decision in stage 1 of the experiment, or let a computer choose for him or her. The stage 1 task consists of picking a number between -6 and 6. If participant A chooses to have the computer make their decision, the computer generates a random number between 0-99 and then makes a decision according to the following table:

Realized number	A's move *a*	Percent
0–6	−6	7
7–8	−5	2
9–15	−4	7
16–19	−3	4
20–21	−2	2
22–26	−1	5
27–39	0	13
40–46	1	7
47–55	2	9
56–62	3	7
63–73	4	11
74–75	5	2
76–99	6	24

Stage 1:

At this stage participant A will either make a decision or let the computer make it for him or her. The following explanation is written as if participant A has decided to make his or her own decision, but the payoff calculations are the same in either case. Participant A can

a) either **increase** your points,
b) or **decrease** your points,
c) or leave your points **unchanged**.

Case a)
Participant A may **increase** your points. In doing so participant A will have to bear **costs**. If participant A spends **one** point, your points increase by **three** points. Suppose for example, that participant A spends three points, then his own points decrease by three points, whereas your points increase by 9 points.

Case b)
Participant A also has the possibility to **decrease** your points in order to increase his own points. Each point that participant A transfers from your points to himself decreases your points by **one** point and increases his own points by **one** point. If, for instance, participant A decreases your points by 3 points, his own points increase by 3 points.

Case c)
Finally, participant A also has the possibility to leave your points **unchanged**. In this case his points will remain unchanged, too.

All three possibilities are shown in the table below: The column "participant A's decision" shows all **possible decisions** for participant A. Please note that an **increase** of your points is indicated by **positive numbers**, whereas a **decrease** of your points is indicated by **negative numbers**. The maximum by which your points can be decreased is 6 points. This means that participant A may choose a number between **-6** and **-1**. If participant A wishes to increase your points, he chooses a number between **+1** and **+6**. If he wishes to leave your points unchanged, he chooses **0**.

The last two columns indicate the points at the end of stage 1, which result from different decisions.

	Participant A's decision	Participant A's points at the end of stage 1	Your points at the end of stage 1
A increases your points to his favor (Case a)	**+6**	12 - 6 = **6**	12 + 18 = **30**
	+5	12 - 5 = **7**	12 + 15 = **27**
	+4	12 - 4 = **8**	12 + 12 = **24**
	+3	12 - 3 = **9**	12 + 9 = **21**
	+2	12 - 2 = **10**	12 + 6 = **18**
	+1	12 - 1 = **11**	12 + 3 = **15**
(Case c)	**0**	12 + 0 = **12**	12 + 0 = **12**
A decreases your points at his cost (Case b)	**-1**	12 + 1 = **13**	12 - 1 = **11**
	-2	12 + 2 = **14**	12 - 2 = **10**
	-3	12 + 3 = **15**	12 - 3 = **9**
	-4	12 + 4 = **16**	12 - 4 = **8**
	-5	12 + 5 = **17**	12 - 5 = **7**
	-6	12 + 6 = **18**	12 - 6 = **6**

Stage 2:

After participant A has made his decision at stage 1, both the points of A and B have **changed**. (The only instance where this does not hold is if A leaves both points unchanged, i.e., in case c). At the second stage it is your turn, as participant B, to make decisions. Starting from the changed points you can

a) either **increase** the points of A that exists at the end of stage 1,
b) or **decrease** the points of A that exists at the end of stage 1,
b) or leave both points, as they exist at the end of stage 1, **unchanged**.

Case a)
You may **increase** the points of A that exist at the end of stage 1. This is **costly** for you. Each point by which participant A's points is increased will reduce your own points by one **point**. Suppose, e.g., you increase A's points by 3 points; then your own points are also reduced by 3 points.

Case b)
You also have the possibility to **reduce** the points of A that exists at the end of stage 1. This is also **costly** for you. Each point you spend reduces A's points by **three** points. Thus, if you spend, e.g., 3 points, your own points decrease by 3 points and A's points decrease by 9 points.

Case c)
You finally have the possibility to leave A's points available at the end of stage 1 **unchanged**. In this case your own points from stage 1 remains unchanged too.

The following table summarizes all **possible** decisions you can make. The column "your decision as participant B" shows all possible decisions. An **increase** of A's points is achieved by the choice of a **positive** number, a **decrease** by the choice of a **negative** number. Please note that an increase of A's points is possible up to a maximum of 18 points. This means that you may choose a number between **+1** and **+18**. If you wish to decrease A's points, choose a number between **-1** and **-6**. If you wish to leave the points unchanged, choose **0**.

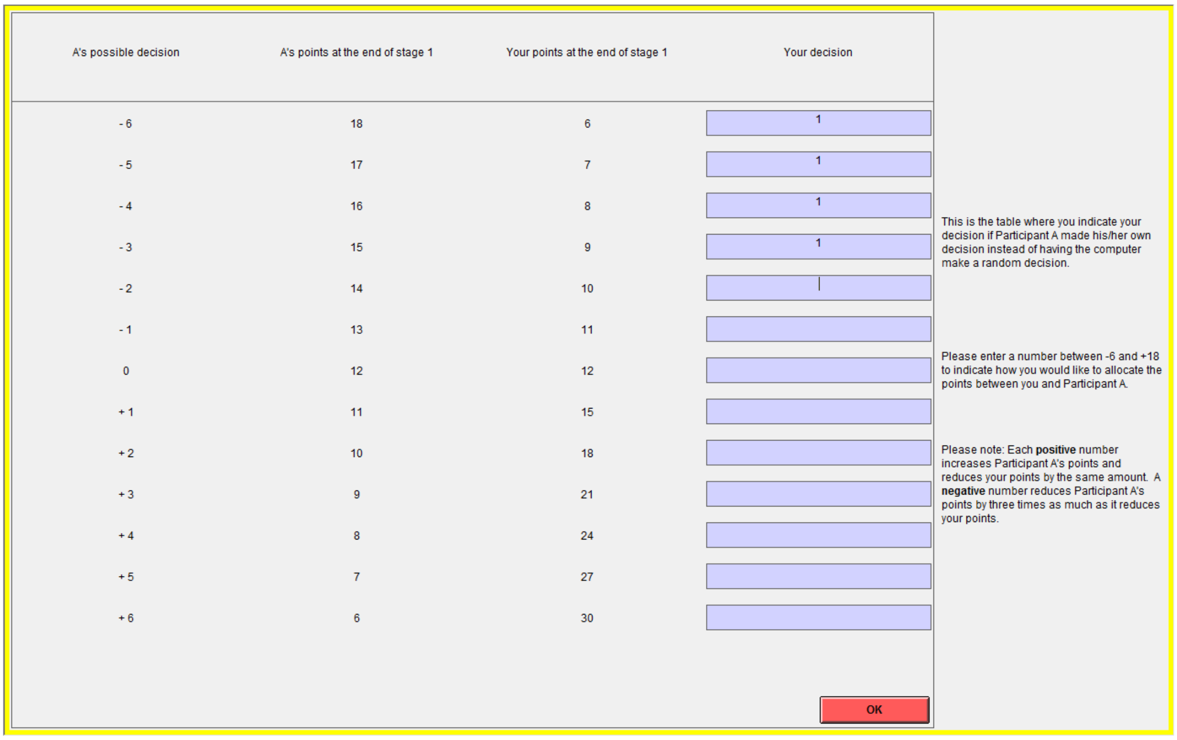
As soon as your decisions are made at stage 2, the **final** points are determined. The final points result from the points at the end of stage 1 and from your decisions as participant B. The table shows these final points of participant A and participant B. The symbol "X" represents the points at the end of stage 1.

	Your decision as participant B	**A's final points** (X = points at the end of stage 1)	**your final points as B** (X = points at the end of stage 1)
	+18	X + 18	X - 18
	+17	X + 17	X - 17
	+16	X + 16	X - 16
	+15	X + 15	X -15
	+14	X + 14	X - 14
	+13	X + 13	X - 13
you increase	**+12**	X + 12	X - 12
A's points	**+11**	X + 11	X - 11
available	**+10**	X + 10	X - 10
at the end	**+9**	X + 9	X - 9
of stage 1	**+8**	X + 8	X - 8
	+7	X + 7	X - 7
(Case a)	**+6**	X + 6	X - 6
	+5	X + 5	X - 5
	+4	X + 4	X - 4
	+3	X + 3	X - 3
	+2	X + 2	X - 2
	+1	X + 1	X - 1
Case c)	**0**	X + 0	X - 0
you decrease	**-1**	X - 3	X - 1
A's points	**-2**	X - 6	X - 2
available	**-3**	X - 9	X - 3
at the end of	**-4**	X - 12	X - 4
stage 1	**-5**	X - 15	X - 5
(Case b)	**-6**	X - 18	X - 6

How do you make your decision?

You will not know whether Participant A will make a decision for him or herself or whether the decision will be made randomly when you make your decision. You also will not know what the decision is. For that reason, you need to make decisions for all possible cases that could occur.

Decisions are made on the computer. There will be two screens – first one for the cases where Participant A made her or her own decision then one where the decision was made randomly. The screen where you insert your decision at stage 2 looks as follows (this is the screen you would see in the case where Participant A made his or her own decision).:



A's possible decision	A's points at the end of stage 1	Your points at the end of stage 1	Your decision
- 6	18	6	1
- 5	17	7	1
- 4	16	8	1
- 3	15	9	1
- 2	14	10	
- 1	13	11	
0	12	12	
+ 1	11	15	
+ 2	10	18	
+ 3	9	21	
+ 4	8	24	
+ 5	7	27	
+ 6	6	30	

This is the table where you indicate your decision if Participant A made his/her own decision instead of having the computer make a random decision.

Please enter a number between -6 and +18 to indicate how you would like to allocate the points between you and Participant A.

Please note: Each **positive** number increases Participant A's points and reduces your points by the same amount. A **negative** number reduces Participant A's points by three times as much as it reduces your points.

OK

In the first column "A's possible decision" you see all possible decisions between **-6** and **+6** participant A can make. In columns two and three ("A's points at the end of the first stage" and "Your points at the end of the first stage") you see which points at the end of the first stage result from these possible decisions for you and participant A. In the fourth column you insert **your decision**. Please indicate, for each of participant A's possible decisions, which decision you make. Thus you must insert 13 numbers. In each row you can insert all numbers between +18 and –6. A positive number increases the points of participant A and decreases your points at the same amount. A negative number decreases the points of participant A – three times as much as it decreases your points. When you have inserted all 13 numbers please press "OK". Then you will go on to the second screen for the "computer decision" case. When you have inserted all 13 of those numbers please press "OK".

Only after you have made your decision, the computer informs you which decision "your" Participant A has made. The final points for you and "your" participant A is determined by the decision of participant A and the number you have inserted in the corresponding row. If, for example, participant A has chosen +3 and you have inserted +4 in the corresponding row, the final split is

for you: 12 + (3 * 3) – 4 = 17
for participant A: 12 – 3 + 4 = 13

At the end of the experiment all participants will learn their counterpart's decision and their total earnings in the experiment.

The experimenters have previously carried out the experiment in which you are now participating. Participant A's in this former experiment have made certain decisions. The following table shows the percentage of the respective decisions made by participants A **in a previous experiment**. Because we have changed the experiment somewhat the results of this experiment are likely to be different.

Decisions of Participant A's in a previous experiment:

Decision	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	**+4**	**+5**	**+6**
Percentage of Participant A's who have chosen the corresponding decision	7 %	2 %	7 %	4 %	2 %	5 %	13 %	7 %	9 %	7 %	11 %	2 %	24 %

Do you have any questions? If you do please raise your hand and someone will be around to answer them. If not, please go ahead and answer the questions in the short quiz on the next two pages.

Quiz

Please calculate the following exercises before the experiment starts. Wrong answers have no consequences for you. The experiment will begin as soon as all participants have correctly solved the problems. When you are done, please consider which decisions you wish to make during the experiment.

1. Stage 1: participant A chooses **+5**.
   Stage 2: participant B chooses **+18**.

   What is participant A's point balance at the end of stage 1? ....
   What is participant B's point balance at the end of stage 1? ....

   What is participant A's final number of points? ....
   What is participant B's final number of points? ....

2. Stage 1: participant A chooses **0**.
   Stage 2: participant B chooses **0**.

   What is participant A's point balance at the end of stage 1? ....
   What is participant B's point balance at the end of stage 1? ....

   What is participant A's final number of points? ....
   What is participant B's final number of points? ....

3. Stage 1: participant A chooses **-6**.
   Stage 2: participant B chooses **-6**.

   What is participant A's point balance at the end of stage 1? ....
   What is participant B's point balance at the end of stage 1? ....

   What is participant A's final number of points? ....
   What is participant B's final number of points? ....

4. Stage 1: participant A chooses **-3**.
   Stage 2: participant B chooses **-1**.

   What is participant A's point balance at the end of stage 1? ....
   What is participant B's point balance at the end of stage 1? ....

   What is participant A's final number of points? ....
   What is participant B's final number of points? ....

5. Stage 1: participant A chooses **-3**.
Stage 2: participant B chooses **+1**.

What is participant A's point balance at the end of stage 1? ....
What is participant B's point balance at the end of stage 1? ....

What is participant A's final number of points? ....
What is participant B's final number of points? ....

6. Stage 1: participant A chooses **+4**.
Stage 2: participant B chooses **-2**.

What is participant A's point balance at the end of stage 1? ....
What is participant B's point balance at the end of stage 1? ....

What is participant A's final number of points? ....
What is participant B's final number of points? ....