EXPERIMENTS ON RECIPROCITY, SOCIAL COMPARISONS AND OVERCONFIDENCE

A thesis submitted in partial fulfilment of the requirements for the

Degree

of Doctor of Philosophy in Economics

in the University of Canterbury

by Katarína Danková

University of Canterbury

2015

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I wish to express my sincere gratitude to my supervisor prof. Maroš Servátka for sharing his knowledge and time and for introducing me to the fields of Behavioural and Experimental Economics. His enthusiastic approach, endless encouragements, excellent problem solving skills and rapid email responses helped me greatly throughout my research journey.

Thank you to the University of Canterbury for the UC Doctoral Scholarship and the Economics and Finance Department for conference and research funding. I would like to thank Jeremy Clark, Miloš Fišar, Bob Reed, John Spraggon, Daniel Schunk, and Daniel Woods for their valuable insights, all the help, feedback and support.

I wish to thank the participants of the 2015 Economics Science Association Meeting in Sydney, the 2014 Ronald Coase Institute Workshop in Manila, the 2014 Australia/New Zealand Workshop on Experimental Economics in Adelaide, the 2014 Economic Science Association Asia-Pacific Meeting in Auckland, the 2013 New Zealand Association of Economists conference in Wellington for helpful comments and suggestions.

Last but not least, I would like to thank my family for all their love, support and understanding.

Abstract

This thesis consists of three chapters focusing on negative reciprocity, transparency and job assignment, and overconfidence. To test my hypotheses I use experimental economics methods, which offer control over the data generating process by motivating people financially. Experimental methods are used to identify causal processes and motivations that can be confounded in field settings. Economics experiments reduce response noise, e.g. extreme outliers probably caused by thoughtless, unmotivated subjects (Smith and Walker, 1993). The central theme of my first two experiments is reciprocity – a kind reaction to kind and generous behaviour and retaliation to hostile behaviour. The first chapter discusses reciprocity in connection with violation of the property rights and the source of the endowment on an individual level. In the second chapter the reciprocity is studied in an employer-worker setting, in the presence of social comparisons. The last chapter deals with firm's overconfidence when deciding about entering the market. All of the chapters study aspects of economic behaviour in social contexts, which have implications for the field. Each of the chapters is described briefly below.

My first chapter experimentally explores the impact of the strength of property rights on retaliation decisions. I induce strong property rights by having experimental subjects earn money by performing a real effort task and weak property rights by endowing them with windfall gains. I ask whether people are less likely to respond to a hostile behaviour with retaliation when earned money as opposed to windfall money is at stake. My experimental design identifies two reasons why property rights might influence the size and frequency of retaliation. The first reason is that retaliation might be

perceived to be more costly when using earned as opposed to windfall money to pay for retaliation. The second reason is related to the violation of property rights. If another person decreases a decision-maker's endowment and the endowment consists of earned money rather than windfall money, the decision-maker might consider it to be a stronger violation of his property rights, which in turn could trigger stronger retaliation. The purpose of this experiment is to separate these two effects. While I find support for the fact that subjects retaliate more because of the violation of their property rights, I also find that participants actually retaliate more with their earned money than with windfall. This suggests that participants do not perceive such retaliation to be more costly but rather that their behaviour is driven by violation of property rights.

The second chapter focuses on the fairness perceptions of the job assignment process in an employer-worker relationship. In reality, employers have at their disposal jobs of different importance, which have to be assigned to different workers. Workers in more important jobs usually get offered higher wages and workers in less important jobs get offered lower wages. If the interpersonal concerns were absent, the employer would provide a higher wage to the worker in the more important job. When a worker decides what wage to accept, he may compare his wage to the other workers' wage. An employer anticipating this might adjust the wage policy in order to avoid unnecessary losses or to maximise profits. I experimentally study the fairness perceptions from the workers' and the employer's point of view. I ask the following questions: 1. Do workers react to the wages paid to their co-workers and does the job assignment procedure affect workers' wage rejections? 2. Do employers react to the fact that workers compare themselves with their co-workers and do they compress wages when the job assignment procedure is perceived less fair? These questions have implications for labour market, in particular a firm's wage policy. If an unfair assignment elicits more wage rejections (i.e. zero profit for the employer and zero wage for the worker) due to social comparisons, it can have detrimental effects on the performance of the firm and thus firms might choose to practice wages secrecy. I examine the impact of job assignment in the presence of social comparisons. In order to test for social comparisons, the worker is only informed about the wage that the employer offered to his co-worker and which job he has been assigned to. Only after workers state their minimum acceptable offers, which determine if the offered wage is accepted or rejected, they come to know their own wage. I posit that a worker's reaction to a specific job assignment depends on the procedure by which they are allocated to the jobs. An assignment to a less important job will not be perceived as unfair if it arises from an unbiased procedure, for example random assignment with equal probabilities. It will, however, be perceived as unfair if workers think that the employer favours some workers over others for reasons that are unrelated to efficiency concerns. This experiment does not provide evidence on social comparisons or employers compressing wages when the assignment to jobs is perceived unfair.

My third chapter is a replication of the high-impact overconfidence and excess entry experiment by Camerer and Lovallo (1999). The topic of overconfidence is crucial for understanding business failures. Camerer and Lovallo were first to directly test overconfidence by measuring economic decisions and personal overconfidence at the same time. Camerer and Lovallo test whether managers' overconfidence about their skills could predictably influence economic behaviour when entering into markets. I implemented Camerer and Lovallo's experiment with modifications reflecting the technological progress of economic experiments of past 15 years. While Camerer and Lovallo run their experiment with male participants (who have been shown to be more overconfident than females), my experiment studies the effect of overconfidence of both genders, making it a more conservative test.

The thesis employs cutting edge techniques from Experimental Economics to study economic decision-making. My research provides empirical evidence on violation of property rights, fairness considerations in labour markets and impact of overconfidence on market entry decisions.

Deputy Vice-Chancellor's Office
Postgraduate Office

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be included for each co-authored work that is included in the thesis. Completed forms should be included at the front (after the thesis abstract) of each copy of the thesis submitted for examination and library deposit.

Please indicate the chapter/section/pages of this thesis that are extracted from co-authored work and provide details of the publication or submission from the extract comes:

Danková, K., & Servátka, M. (2015). The house money effect and negative reciprocity. *Journal of Economic Psychology*, *48*, 60-71 is based on Chapter 1 of this thesis.

Please detail the nature and extent (%) of contribution by the candidate:

Although I am a co-author on the journal article, Katarína has done most of the work on the project, including programming the software, running the experiment and analyzing data, whereas I served as a consultant who helped Katarina develop the initial idea and later turn her dissertation chapter into a polished paper that could be submitted into a journal. Katarína has done about 75% of the work and is the first author on the published article. The chapter features the original (extended) version of the paper.

Certification by Co-authors:

If there is more than one co-author then a single co-author can sign on behalf of all
The undersigned certifys that:

- The above statement correctly reflects the nature and extent of the PhD candidate's contribution to this co-authored work
- In cases where the candidate was the lead author of the co-authored work he or she wrote the text

Name: *Maroš Servátka*	Signature: *Maroš Servátka*	Date: *28.7.2015*

Chapter 1

1 Negative Reciprocity and the House Money Effect

1.1 Introduction

Over the past two decades experimental economics research has provided ample evidence that people care not only about their own material payoffs but also that they are willing to forego significant income in order to influence the payoffs of others (e.g. Cox and Deck, 2005; Cox, Sadiraj, and Sadiraj, 2008; Falk, Fehr, and Fischbacher, 2003; Güth, Schmittberger, and Schwarze, 1982). Much attention has been devoted particularly to reciprocity – a tendency to react to the kind actions of others with kind responses and to the hostile actions of others with retaliation.[1] In the vast majority of laboratory experiments documenting the existence of reciprocity subjects are endowed with start-up windfall. This endowment serves as starting capital from which subjects draw when making decisions of interest to the experimenter, potentially creating a "house money effect." The current chapter explores the implications of endowing subjects with windfall gains on their observed reciprocal behaviour. In particular, I study (1) whether earned income makes (negative) reciprocity more costly and thus leads to less retaliation than windfall and (2) whether appropriating another's earned money is considered a stronger violation of property rights than appropriating windfall money and thus results in more retaliation.

[1] For example Ostrom, Walker, and Gardner (1992), Fehr and Gächter (2000), Ostrom and Walker (2005); see also Camerer (2003), Cox (2013) for surveys and (Rabin, 1993), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006), Cox, Friedman, and Gjerstad (2007) and Cox, Friedman, and Sadiraj (2008) for theoretical approaches to modeling reciprocity. Also, while not a model of reciprocity per se, inequity-aversion theory (Fehr and Schmidt, 1999) is sometimes used to explain behaviour of fair-minded people who want to achieve more equitable final outcomes.

Since money is fungible, why should the source of endowment matter? According to mental accounting (Thaler, 1985), different sources of income might lead to different ways of spending. If the costs of obtaining an endowment vary, people may place such endowments in separate mental accounts, which might in turn lead to different choices. In particular, having to earn an endowment could increase the perceived cost of reciprocation and thus diminish its frequency and/or extent (Anderson and Putterman, 2006; Carpenter, 2007). This conjecture is consistent with the results of previous studies, which show that having subjects earn money in another task prior to making decisions produced more self-serving behaviour both in the laboratory (for example, Bosman, Sutter, and van Winden, 2005; Cherry, Frykblom, and Shogren, 2002; Oxoby and Spraggon, 2008) and in the field (Clingingsmith, 2013). Since reciprocation is costly, it is possible that the experiments in which subjects use windfall money to pay for retaliation overestimate the extent of reciprocal behaviour. Previous research conversely shows that generosity observed under laboratory conditions is often greater than one would observe in the field and that this might be partly due to the house money effect (Carlsson, He, and Martinsson, 2013). However, the existing literature does not permit a conclusion whether reciprocity observed in laboratory conditions is subject to the house money effect as well as this had not been tested previously.

As highlighted above, there are two reasons why negative reciprocity might be sensitive to the source of the endowment. The first is that if money used to pay for retaliation is earned as opposed to received from the experimenter, one might perceive the retaliation to be more costly due to this money being in a different mental account. The second reason is related to property rights. If another person decreased the decision-maker's endowment consisting of earned money rather than windfall money, the decision-maker might consider it to be a stronger violation of his property rights, which in turn could trigger stronger retaliation. To tackle the implications of the house money effect for reciprocal behaviour I present an experiment in which the source of the endowment varies. However, rather than focusing on the overall house money effect as most of the previous literature, the current experimental design allows to identify two reasons why endowing subjects with a windfall might impact the observed level of negative reciprocity.

1.2 Literature Review

Subjects in the current experiment interact in a two-player Taking Game in which the First Mover (henceforth FM) decides whether or not to take a sum of money from the Second Mover's (henceforth SM) endowment. If the FM decides not to take, the game ends and both players keep their initial endowments. If the FM takes money from the SM, the SM can retaliate in return. To distinguish between the reasons that could lead to different retaliatory behaviour I implement a 2x2 experimental design, in which I vary whether the FM takes the earned or windfall part of the SM's endowment and whether the SM then retaliates with his earned money or with windfall. In the first treatment, the SM's endowment consists entirely of windfall received from the experimenter. Whatever is left of the endowment after the FM's decision, the SM can use to purchase retaliation. In the second and third treatment, the SM's endowment consists of funds earned in a real-effort task as well as of windfall. In one case, the FM takes the windfall part of the endowment and the SM can retaliate using his earned money whereas in the other case the FM takes the earned part and the SM can retaliate with windfall. In the fourth treatment, the SM's endowment consists entirely of earned money.

This study builds on earlier work of three types: research on negative reciprocity, the house money effect and mental accounting. I discuss each of them in some detail in relation to the research question.

Reciprocity plays an important role in labour relationships as it has a potential to increase efficiency through enforcement of incomplete contracts. For example, Fehr, Gächter, and Kirchsteiger (1997) experimentally investigate whether reciprocity motives are sufficiently strong to overcome contract enforcement problems. In a series of market experiments they isolate reciprocity effects on contract enforcement. Somebody who offers a generous contract, which the trading partner subsequently violated, might be willing to punish the trading partner. Yet, if the trading partner anticipates this willingness to punish, he has a reason not to violate the contract in the first instance. Punishing unfair behaviour is also a form of negatively reciprocal behaviour, i.e. people retaliate towards those who were unkind to them or even hurt them. Fehr, Gächter, and Kirchsteiger (1997) implement three experimental conditions. In the first treatment, labelled the no-reciprocity treatment, contract terms are exogenously enforced so that reciprocity cannot contribute to contract enforcement. Firms post employment contracts in a competitive market and once the contract is accepted the worker has to supply the effort level, which is fixed by the experimenter. In the second treatment, labelled the weak reciprocity treatment, only one side of the market can respond reciprocally to the previous action of the trading partner. In this treatment, firms post employment contracts and once the contract is accepted the worker has to decide whether to supply effort that is demanded by the firm or whether to shirk. And lastly, in the third treatment, labelled the strong reciprocity treatment, after workers' effort decision firms have the opportunity to punish or reward the worker. Both reactions are costly to the firm. In all treatments, there is an exogenous excess supply of workers. Their results show that in the weak reciprocity treatment the strength of workers' reciprocal responses depends on the details of the pecuniary incentives.[2] In the strong reciprocity treatment, reciprocity is a powerful device for the enforcement of contracts, i.e. firms reciprocate, workers anticipate this and shirk less than in the weak reciprocity treatment and firms demand and enforce effort levels above incentive compatible levels.

The notion of loyalty follows from reciprocity. Firms value loyal workers who are committed to the goals of the firm. Loyalty means that workers take into account the interests of their employers. If employers also take into account the interests of their workers, a positive valuation of the employer's payoff can be created (Fehr and Falk, 2002). In the same vein, workers have many opportunities to take advantage of employers. Poor treatment of workers could lead to negatively reciprocal behaviour such as low effort or even sabotage. Firms (and their managers) are well aware of potential repercussions from diminished morale and loyalty and try to circumvent them by implementing sensible strategies, such as not lowering wages following a demand shock, leading to downward wage rigidity (Bewley, 1999). A gift-exchange game has been the workhorse used to study various aspects of labour market relationships and incomplete contracts (see Charness and Kuhn, 2011 for a nice survey). While most laboratory gift-exchange experiments induce costs of effort using the house money approach, there exist

[2] An additional treatment where losses were possible was run in order to determine the sensitivity of reciprocal behaviour. In the original weak reciprocity treatment workers almost always choose the lowest effort level if they shirk, whereas in the additional weak reciprocity treatment partial shirking is frequent.

some experiments (both lab and field) that involve a real-effort task (e.g. Gneezy and List, 2006; Kube, Maréchal, and Puppe, 2012) and thus circumvent the potential house money effect.

For example, Gneezy and List (2006) focus on real effort in labour markets using two field experiments. In the first one, they invited people who were not informed that they were taking part in an experiment, to computerize the holdings of a small library at a university. The first treatment paid a flat wage of $12 per hour, as promised. In the second treatment, once the task was explained to the participants, the participants were told that they would be paid $20 per hour, not the $12 that had been promised. The findings are in line with the gift exchange hypothesis, participants in the $20 treatment provided significantly higher effort in the first 90 minutes than participants in the $12 treatment. After 90 minutes on the job, however, effort levels were the same across the two treatments. In their second field experiment the authors invited students to take part in a door-to-door fundraising drive to support a research centre at a university. The participants were told that they would be paid $10 per hour. An important difference from the previous experiment is that workers have a better idea about the surplus created, i.e. how many and how large the contributions were, and how much the employer valued the task. This is an important feature of the second field experiment, because if the contributions are known, the share of the contributions that the workers receive will determine whether they perceive their wage as fair. Alternatively, if workers know only the promised wage and not the contributions (how much the task is worth to the employer), as in the library task, only the promised wage can serve as a reference point. The first treatment was a flat wage of $10 per hour, as promised. In the second treatment, students were told that they would be paid $20 per hour, not the $10 that had been advertised. The results are that students in the $20 treatment raised significantly more money in the first few hours of the task than students in the $10 treatment, but after a few hours the observed outcomes were the same.

Also, Kube, Maréchal, and Puppe (2012) conducted a controlled field experiment to measure the extent to which monetary and non-monetary gifts affect workers' performance. They recruited workers to catalogue books from a library. The job was announced with an hourly wage of 12 EUR. In the Baseline treatment the workers were paid 12 EUR. In the Money treatment, the workers received a monetary gift in the form of a 20% wage increase (7 EUR extra). In the Bottle treatment, instead of the pay raise, workers received a thermos bottle worth 7 EUR on the top of their wage. The results show that the nature of gifts determines the prevalence and strength of reciprocal behaviour. The cash gift (Money treatment) had no statistically significant impact on workers' productivity compared to the Baseline. The bottle, however, resulted on average in a 25% higher work performance. They replicated the results from the Bottle treatment with a Price Tag treatment where they explicitly mentioned the bottle's market price. Workers produced almost an equal output in the Price Tag and Bottle treatments. An additional treatment Choice was conducted, where the workers could choose between receiving cash or the bottle. More than 80% of workers opted for the cash gift, but workers' output was statistically significantly higher in the Choice treatment than in the Baseline and Money treatments. To provide a more direct test of whether time and effort matter for workers' output, they conducted an Origami treatment. In this treatment the employer gave the workers money in the form of an origami (i.e. the money was artistically folded and wrapped). The origami was identical to the cash gift, except that the employer had invested more time and effort into the gift. The statistically significant results show that workers reciprocated the Origami cash gift by producing 30% more output relative to the Baseline and 23% more than in treatment Money. Non-monetary gifts have thus a much stronger impact than monetary gifts of equivalent value and when offered a choice, workers prefer receiving the money but reciprocate as if they received a non-monetary gift. Although these experiments involve a real effort task, I am unaware of any studies that compare behaviour in a gift-exchange game with windfall versus earned endowments.

Another strand of literature on (negative) reciprocity studies the impact of punishment or sanctions on social norms. In a seminal paper, Ostrom, Walker, and Gardner (1992) show that introducing costly punishment in a common pool resource game can overcome strong self-interest of individual appropriators and lead to a mutually efficient outcome. Fehr and Gächter (2000) explain that many co-operators have an aversion to being exploited and are willing to punish free-riders in the voluntary contribution mechanism (e.g. Isaac, McCue, and Plott, 1985). These two papers started a whole new area of research dealing with various aspects of punishment (e.g. demand for punishment in Anderson and Putterman, 2006; Carpenter, 2007; fueds in Nikiforakis and Engelmann, 2011; punishment technology in Nikiforakis and Normann, 2008), anomalies such as counter-punishment (e.g. Nikiforakis, 2008) and anti-social punishment (e.g. Herrmann, Thöni, and Gächter, 2008).

Results from several extensive-form game experiments also show that subjects frequently exercise the explicit or implicit option to punish non-cooperation or unfair behaviour (moonlighting game in Abbink, Irlenbusch, and Renner, 2000; prisoner's dilemma in Clark and Sefton, 2001; Cox and Deck, 2005; moonlighting game in Cox, Sadiraj, and Sadiraj, 2008; ultimatum game in Falk, Fehr, and Fischbacher, 2003; and Güth, Schmittberger, and Schwarze, 1982). Even if punishment is costly to subjects, those who are perceived to be unkind or reveal malevolent or selfish intentions are often punished (Bosman and van Winden, 2002; Pereira and Silva, 2006).

Experimental literature thus provides ample evidence of reciprocity under controlled laboratory conditions. In most of these experiments participants are endowed with start-up funds from the experimenter. Such funds might be treated as a windfall gain and thereby create the house money effect, meaning that people might spend or invest such money more recklessly than they would their own. The house money effect was first evidenced by Thaler and Johnson (1990) in a lottery choice experiment in which losses were subtracted from subjects' initial endowments. Several studies point out that the observed behaviour might differ if subjects receive windfall endowments as opposed to when these endowments are earned. In the following I list some examples of 'real-effort' endowment earning tasks. Cox and Hall (2010) use a computerised "whack-a-mole game" as a real effort task to earn the endowments in their modification of an investment game (Berg, Dickhaut, and McCabe, 1995).[3] Harrison (2007) asked subjects to bring their own money in order to participate in a public goods game; Reinstein and Riener (2012) in their charitable giving experiment (dictator game with a charity acting as a recipient) asked subjects to earn their endowment by adding up five two-digit numbers; (Cherry, Frykblom, and Shogren, 2002; Cox and Hall, 2010; Harrison, 2007; Hoffman and

[3] In the whack-a-mole game there is a 6 by 4 grid of moles and holes on the field. Each time the subject mouse-clicks a mole picture the picture box shows a hole picture. The object of the game is to mouse-click all of the moles until the field is clear of moles.

Spitzer, 1985; Reinstein and Riener, 2012; Rutström and Williams, 2000) elicit experimental subjects' preferences for income redistribution using the task of the Tower of Hanoi puzzle. Cherry, Frykblom, and Shogren (2002) run a series of dictator game experiments in which the endowments were earned through solving GMAT questions. Such a design yielded significantly less generous dictator behaviour than a control treatment where the dictator's endowment was randomly determined. Having to earn the endowment creates a property right entitlement and as a result leads to a more self-regarding behaviour by the person who earned it.[4]

Cox, Servátka, and Vadovič (2014) also have their subjects earn endowments by solving GMAT questions in an experiment testing whether acts of commission that overturn the status quo generate stronger reciprocal responses than acts of omission that uphold it. Their design consists of two treatments that vary with respect to the size of initial endowments, thereby inducing different status quos. In their first treatment the FM can either give a part of his endowment to the SM (which is an act of commission) or do nothing (an act of omission). The SM then decides whether or not to reward the FM for giving him money and whether or not to punish him for not giving. In the second treatment, the FM can either take a part of the SM's endowment (an act of commission) or do nothing (an act of omission). The SM then decides whether or not to punish the FM for taking his money and whether or not to reward him for not taking. Cox, Servátka, and Vadovič (2014) find that earning endowments significantly affected giving and taking behaviour by the FMs who gave less often to the SMs as well as took less frequently

[4] This is contrasted with other participants respecting the established property rights to one's endowment (Cox, Servátka, and Vadovič, 2014; Hoffman and Spitzer, 1985; Rutström and Williams, 2000). A similar pattern has been observed in experiments where the roles in a game are earned (e.g. Erkal, Gangadharan, and Nikiforakis, 2011; Hoffman et al., 1994).

from them when the endowments were earned, but had an insignificant effect on SMs' reciprocal responses.

Clark (2002) finds no effect of windfall endowments in the voluntary contributions mechanism public goods game using unconditional nonparametric methods. Harrison (2007) shows that the same data display a significant effect when analysing responses at the individual level and accounting for the error structure of the panel data. According to Harrison, there were more free riders in the windfall treatment than in the own money treatment, but the windfall had no clear influence on the levels of positive contributions. Finally, Carlsson, He, and Martinsson (2013) used a 2x2 design combining laboratory and field experiments to examine the impact of windfall money on generosity. In both environments they found that subjects donate more when the endowment is a windfall.

The most closely related studies to ours are by Bosman and van Winden (2002) and Bosman, Sutter, and van Winden (2005) who respectively examine the impact of emotions and real effort on behaviour in the Power-to-Take Game. In this two-player sequential-moves game one player (the take authority) can claim any fraction $t \in [0,1]$ of the endowment of the other player (the responder). The latter has an opportunity to diminish the claim by choosing a destruction rate $d \in [0,1]$ and destroying his own endowment. This results in the payoff $E_{Take} + t(1-d)E_{Resp}$ for the take authority and $(1-t)(1-d)E_{Resp}$ for the responder. In their experiment both players earn their endowments. In such an environment the intensity of negative emotions experienced by the responder is positively correlated with the taking rate and the probability of endowment destruction is positively correlated with both the intensity of experienced negative emotions and the taking rate.

In a follow up paper Bosman, Sutter, and van Winden (2005) introduce a treatment, in which subjects are endowed with windfall money, to explore whether the behaviour of players is influenced by the fact that they have to earn their endowments. The results show that the taking behaviour does not depend on real effort, i.e. earning the endowment, and that responders destroy their endowment more often and in greater amount on aggregate with windfall endowment than with earned endowment, providing evidence of the house money effect. Also, without earning their own endowments, intermediate amounts of destruction are chosen more often whereas if effort is provided, the destruction is all or nothing. In our experiment I also study how earning endowments in a real-effort task influences the decision making of subjects, however, the main contribution lies in decomposition of the house money effect.

From some of the previous studies it appears that if endowments are earned, subjects act in a more self-regarding manner. What causes such changes in behaviour? Mental accounting may shed some light on the issue. According to the principle of fungibility all money is the same regardless of its origin or intended use. However, people often treat money differently depending on its source and separate their funds into mental accounts based on subjective criteria (Thaler, 1985). Mental accounting is a cognitive process by which people keep track of the flow of their money and keep their spending under control. Thaler (1985) distinguished among three components of mental accounting. The first one captures how outcomes are perceived and experienced, and how decisions are made and subsequently evaluated. The second component involves the assignment of activities to specific accounts. Both the sources and uses of funds are labelled in real as well as in mental accounting systems. The third component of mental accounting concerns the frequency with which accounts are evaluated. Thus according to mental accounting, money in one mental account is not a perfect substitute for money in another account, which might lead to violations of the normative economic principle of fungibility. Different sources of experimental endowment might therefore have implications for observed subject behaviour.

1.3 Experimental Design And Procedures

The purpose of the current experiment is to test whether negative reciprocity is subject to the house money effect and to identify the driving reasons behind it as receiving a windfall endowment from the experimenter might impact the size and/or frequency of retaliation. The existing experimental economics literature describes (at least) three methods of controlling for the house money effect. Clark (2002) implemented a control for the house money effect by having subjects bring their own money to the experiment. While this is certainly a possibility, most experimental economics laboratories, including ours, advertise that subjects will on average earn a significant amount of cash from participation. Since in the current design – as described below – it is possible to suffer a loss, such a method could impact the reputation and credibility of the laboratory and discourage future participation.

A second method, used by Cárdenas et al. (2013), involves giving money to subjects in advance and requires them to bring it to the experiment. Such a method enables the participants to "bond" with the funds as after a period of time of being in possession they might start considering them their own. However, this method bears a risk that subjects will not show up for the experiment or they will not bring the money with them.

A third method involves creating property right entitlements towards the initial endowments by having subjects earn them in a real-effort task as in Hoffman and Spitzer (1985) and Cherry, Frykblom, and Shogren (2002). Such a procedure serves as an analogue to everyday life where people exchange their time and effort for monetary payments. While it is possible that the three methods lead to different levels of entitlement, In my opinion the third method was quite natural while also being the most practical one and decided to implement it in the current experiment.

In my setup, there are two reasons why the source of endowment might matter. The first reason is that if a person had to earn the money he uses to pay for retaliation, as opposed to receiving it from the experimenter, he might perceive the retaliation to be more costly due to placing this money in a different mental account. He would therefore retaliate less. The second reason is that if earned money, as opposed to windfall, is taken from the person, he might consider it to be a stronger violation of his property rights, which in turn could trigger more retaliation.

In order to separate these two effects, I introduce a Taking Game played by two players, the FM and the SM. The experiment consists of four treatments (presented in Figure 1.1): Windfall/Windfall, Windfall/Earned, Earned/Windfall, and Earned/Earned, implemented in an across-subjects design. The treatments differ in the source of the SM's endowment, which (1) could be taken by the FM (the first label) and (2) is used to retaliate (the second label).

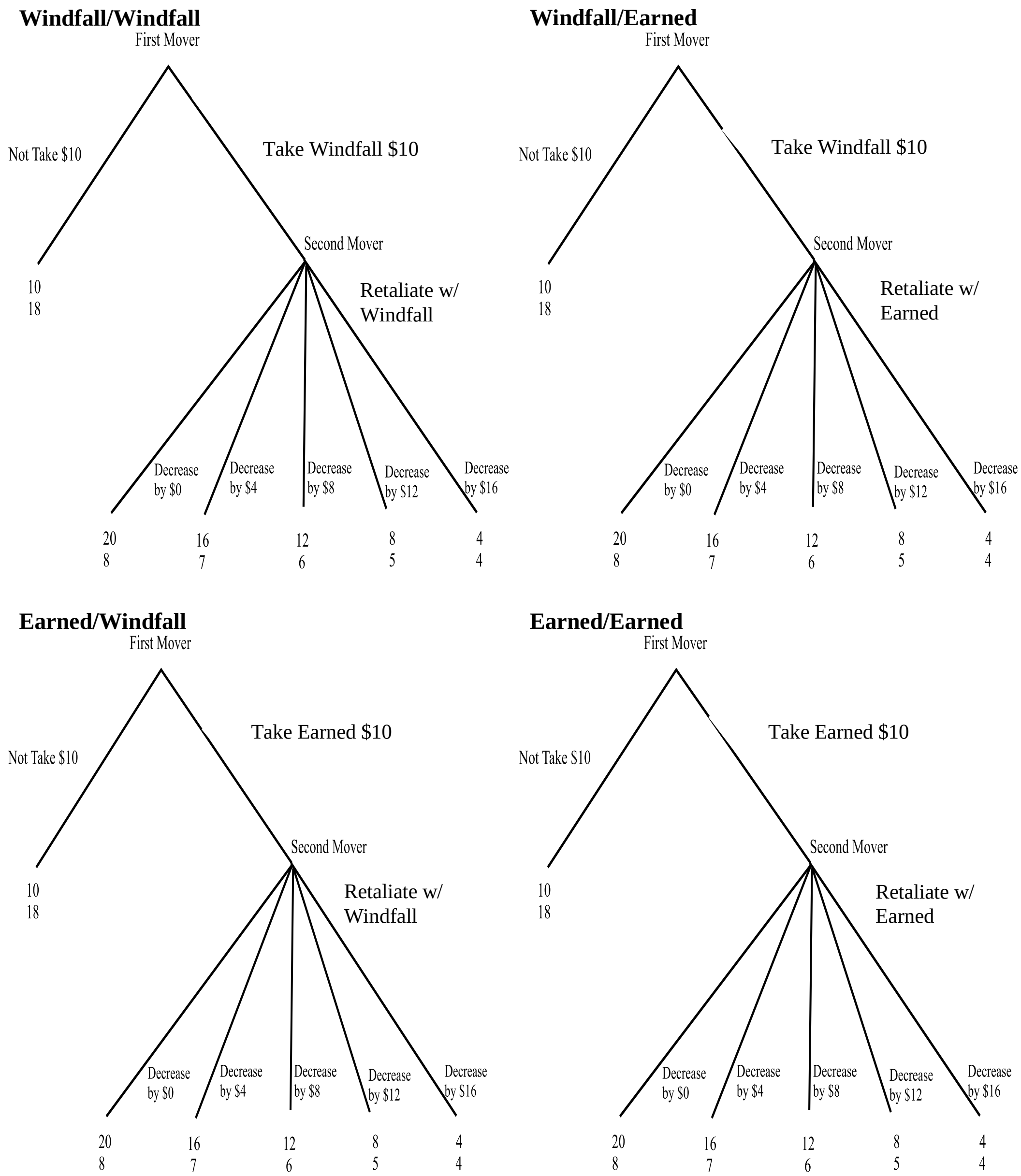
In the Windfall/Windfall treatment the FM receives $10 and the SM receives $18 from the experimenter as initial endowments. There is no real-effort task performed by either player. Both players then proceed to playing the Taking Game, which is described below.

In the Windfall/Earned and Earned/Windfall treatments, the SM's endowment consists of funds earned in a real-effort task as well as of windfall funds. In the Windfall/Earned treatment the FM and the SM are endowed with $10 each. The SM performs a real-effort task, in which he can earn an additional $8. The real-effort task consists of cutting posters inviting students to participate in economics experiments in our laboratory. This particular task does not require any specific skills and I decided to implement it because it represents a meaningful activity as subjects could later see the posters on notice boards around the university campus. I asked SMs to neatly cut the bottom part of 20 posters to create stubs that included a web page link of the database where interested students could register for experiments. For accomplishing the task, each SM earned $8. After the SMs finished the real-effort task, all subjects proceeded to playing the Taking Game.

In the Windfall/Earned treatment the FM can take the windfall part of the endowment (a weak violation of property rights) and the SM can retaliate using his earned money. In the Earned/Windfall treatment this is reversed – the FM can take the earned part of the endowment (a strong violation of property rights) and the SM can retaliate using windfall funds. Thus, in the Earned/Windfall treatment, the FM is endowed with $10 and the SM with $8. On the top of that, the SM can earn additional $10 for cutting posters.

Finally, the $18 endowment of the SM in the Earned/Earned treatment consists entirely of funds earned in the real-effort task. The Taking Game that was played in all four treatments is presented in Figure 1.1.

Figure 1.1 The Taking Game and the Implemented Treatments



The game was played only once thus all decisions are independent. In Stage 1 of the game the FM decides whether to *Take* or *Not Take $10* from the SM. If the FM does *Not Take $10,* this yields a payoff of $10 to the FM and a payoff of $18 to the SM. If the FM *Takes $10* from the SM, the game proceeds with Stage 2, where the SM decides whether to retaliate by decreasing the FM's payoff. Every $1 the SM uses for retaliation reduces the FMs payout by $4. The SM can use up to $4 from his own endowment to decrease the FM's payoff up to $16. If the SM does not wish to retaliate against the FM, he can do so by choosing the *Decrease by $0* option.[5]

For this design it was crucial that subjects recognized that the SM's endowment consisted of two parts – one that could be taken by the FM and the other that could be used to pay for retaliation. In order to highlight this I deliberately chose different (yet relatively close in value) amounts to represent these two parts: in all treatments $10 could be taken and $8 was used for retaliation. What differed between treatments was the source. The instructions were framed in a way to ensure that subjects understood which part of their total endowment was being taken and which was used for retaliation. Such framing might aid creating two separate mental accounts. The current design allows identifying the reasons why the source of endowment might matter for the frequency and/or extent of retaliation. I test the following hypotheses:

[5] My Taking Game differs from Bosman, Sutter, and van Winden (2005) Power-to-Take Game in the following ways. In the Taking Game the FM has only two options, i.e. to Take $10 from the SM or Not Take $10, whereas in the Power-to-Take Game the FM chooses a taking rate from the interval [0, 1]. The action space for the SM differs as well. While in the Taking Game the SM chooses one of five available actions with a destruction technology 1:4, in the Power-to-Take Game he chooses a destruction rate from the [0, 1] interval. The Taking Game is more appropriate for my purposes because it allows me to separate the two reasons behind the house money effect.

Hypothesis 1: There will be less retaliation in the Windfall/Earned treatment than in the Windfall/Windfall treatment (*the cost-of-reciprocity effect under a weak violation of property rights*).

Hypothesis 2: There will be more retaliation in the Earned/Windfall treatment than in the Windfall/Windfall treatment (*the property-rights-violation effect under a low cost of reciprocity*).

By combining Hypothesis 1 and Hypothesis 2 together (conditional on taking), I expect to see more retaliation in the Earned/Windfall than in the Windfall/Earned treatment. This is due to the cost-of-reciprocity effect and the property-rights-violation effect going in the same direction when comparing these two treatments.

Hypothesis 3: There will be more retaliation in the Earned/Windfall treatment than in the Windfall/Earned treatment (*combined cost-of-reciprocity and property-rights-violation effects*).

Hypothesis 4: There will be less retaliation in the Earned/Earned treatment is than in the Earned/Windfall treatment (*the cost-of-reciprocity effect under a strong violation of property rights*).

Hypothesis 5: There will be more retaliation in the Earned/Earned treatment than in the Windfall/Earned treatment (*the property-rights-violation effect under a high cost of reciprocity*).

Note that I do not form an explicit hypothesis about the Windfall/Windfall vs. Earned/Earned treatment comparison due to the cost-of-reciprocity effect and the property-rights-violation effect going in the opposite directions between these two treatments.

A total of 296 subjects participated in the experiment. The experimental sessions were conducted in the New Zealand Experimental Economics Laboratory (NZEEL) at the University of Canterbury. Subjects were recruited using the online database system ORSEE (Greiner, 2004). Each subject only participated in a single session of the study. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

The number of subjects in a session varied from twenty to thirty-six. All sessions were run under a single-blind social distance protocol. On average, a session lasted 60 minutes including the payment and subjects earned 15.55 NZD (New Zealand dollars).[6]

Upon entering the laboratory, subjects were randomly assigned into Group A (FMs) and Group B (SMs) by drawing a letter (A or B) from a manila envelope and asked to sit in a cubicle in an appropriate row. At the beginning of the experiment instructions (provided in Appendix B) were handed out, as well as projected onto a screen. In the Windfall/Earned, Earned/Windfall and Earned/Earned treatment, Task 1 instructions were first handed out to all Group A and Group B persons and read aloud. Group B persons were then given scissors and 20 NZEEL posters and asked to cut them along the perforated lines. Meanwhile, Group A persons were asked to sit quietly and not to use the lab computers. Once all Group B subjects completed the task, the posters and scissors were collected. Task 2 instructions were handed out, projected onto the screen

[6] The minimum hourly wage in New Zealand at the time of the experiment was 13.50 NZD.

and read aloud to ensure common knowledge of the fact that Group B persons had to earn (a part of) their starting balance.[7]

Any questions arising were answered in private. Subjects had to answer all control questions correctly before they could proceed to the decision making part of the experiment, which was run using the strategy method (Selten, 1967).[8] After the control questions, subjects entered their decisions. Upon the completion of the experiment, they were also asked to fill out a questionnaire. Subjects were then called one by one to receive their payment in private in the payment room in the back of the laboratory.

1.4 Results

1.4.1 First Movers' Behaviour

While the main focus of the current experiment is on the behaviour of SMs, I present the results in the order of play. The taking behaviour of self-regarding FMs will in part be indicative of their expectations of SMs' reactions, meaning there should be an inverse relationship between the frequency of taking and the expected retaliation. I test

[7] Since the Windfall/Windfall treatment did not involve cutting posters, Task 2 instructions were simply called Task instructions.

[8] In the strategy method the SM makes conditional decisions for each possible information set, rather than learning about the action of the FM and then choosing a response, as in the direct-response method. In our experiment this amounts to the SM making a retaliation decision without knowing whether the FM decided to take money or not. Brandts and Charness (2011), who survey experiments regarding whether the strategy method leads to different behaviour than the direct-response method, argue that the strategy method may lead subjects to better insights into the motives and thought-processes underlying their decisions through the analysis of a complete strategy. They find that in the majority of surveyed studies there is either no difference in behaviour between the two methods or that the evidence is mixed. Using the strategy method in our experiment could be considered a conservative approach, since if the "hot" direct-response method were to produce more retaliation than the "cold" strategy method I could reasonably expect my results to be even stronger.

predictions about the FMs' behaviour derived from the hypotheses about their SM's reciprocal reactions.[9]

In summary, the hypotheses predict more retaliation in the Earned/Windfall treatment than in any other treatment and more retaliation in both Earned/Earned and Windfall/Windfall treatments than in the Windfall/Earned treatment. This should result in less taking in the Earned/Windfall treatment than in any other treatment and less taking in both Earned/Earned and Windfall/Windfall treatments than in the Windfall/Earned treatment.

This is not what I find in the data as can be seen from Panel A of Table 1.1. The highest frequency of taking occurred in the Windfall/Windfall treatment where 31 out of 34 (91.2%) FMs took $10 from the SMs. In the Earned/Earned treatment 30 out of 36 (83.3%) FMs took $10 from the SMs; in the Windfall/Earned treatment 33 out of 41 (80.5%); and in the Earned/Windfall treatment 28 out of 37 (75.7%).

I test for treatment differences in the frequency of taking behaviour using the Fisher's exact test, reported in Panel B of Table 1.1. None of the comparisons is statistically significant when using a 2-sided test. However, if one were to use a 1-sided test (given that there exist *ex ante* predictions), the frequency of taking behaviour in Earned/Windfall would be mildly statistically significantly lower than in Windfall/Windfall ($p = 0.076$). The probit regression of the FMs' taking behaviour on treatment dummies (presented in Panel C of Table 1.1; the regression coefficients

[9] The taking behaviour of FMs might also depend on whether they consider the SMs' endowments to be fungible. If the FMs expect different reactions from the SMs depending on whether earned or windfall money was taken, it should lead to higher taking rates in the Windfall/Windfall and Windfall/Earned treatments than in the Earned/Windfall and Earned/Earned treatments. The effect of FMs' expectations could be amplified by potentially a different level of property rights – a FM might be willing to take the windfall part of the SM's endowment, but respect the SM's entitlement to the earned part.

represent differences from the Windfall/Windfall treatment) supports the results of the non-parametric tests.

Table 1.1 First Mover Behaviour

Panel A: Frequency of Taking			
Windfall/Windfall	Windfall/Earned	Earned/Windfall	Earned/Earned
31/34 (91.2%)	33/41 (80.5%)	28/37 (75.7%)	30/36 (83.3%)

Panel B: Fisher's Exact Test, P-values (2-sided)[a]			
	Windfall/Windfall	Windfall/Earned	Earned/Windfall
Windfall/Earned	0.326	-	-
Earned/Windfall	0.115	0.784	-
Earned/Earned	0.479	0.777	0.564

Panel C: Probit Regression of the First Mover Taking Behaviour on Treatment Dummies [a]				
		pseudo-R^2	χ^2	Prob> χ^2
		0.024	3.31	0.346
Treatment	Coefficient	Marginal	z	P>z
Windfall/Earned	-0.493	-0.135	-1.30	0.192
Earned/Windfall	-0.656	-0.188	-1.73	0.083
Earned/Earned	-0.384	-0.105	-0.98	0.328
Constant	1.352	0.831	4.45	0.001

[a] Windfall/Windfall is the omitted control variable.

1.4.2 Second Movers' Behaviour

SMs' behaviour is summarised in Table 1.2. The first column reports the number and frequency of SMs who decided to retaliate against their paired FM for taking $10. Recall that I used the strategy method and therefore have observations for all SMs who participated in the experiment. To my surprise, the highest frequency of retaliation occurred in the Windfall/Earned treatment where 51.2% SMs retaliated spending on average $1.76, followed by the Earned/Earned (50% and $1.64), Earned/Windfall (46% and $1.38), and Windfall/Windfall (24.5% and $0.77) treatments. The last column of Table 1.2 presents the consequences of retaliation for the FMs' earnings.

Table 1.2 Second Mover Behaviour

	Number of Retaliating Subjects	**Avg. Amount Spent on Retaliation($)**	**Standard Deviation**	**SMs Decreased FMs' Earnings on Average by ($)**
Windfall/Windfall (n=34)	8/34 (24.5%)	0.77	1.47	3.08
Windfall/Earned (n=41)	21/41 (51.2%)	1.76	1.86	7.04
Earned/Windfall (n=37)	17/37 (46.0%)	1.38	1.69	5.52
Earned/Earned (n=36)	18/36 (50.0%)	1.64	1.81	6.56

The distribution of retaliation across the four treatments is presented in Figure 1.2 and the resulting distribution of final payoffs in Table1.3. The largest difference in SM behaviour appears to be in the $0 and $4 spent on retaliation categories. While in the Windfall/Windfall treatment only 24.5% of SMs retaliated, in treatments where SMs earned at least a part of their endowment, the frequency of retaliation was much higher: 51.2 % in Windfall/Earned, 46% in Earned/Windfall and 50% in Earned/Earned.

Figure 1.2 Distribution of Second Movers' Retaliation Decisions

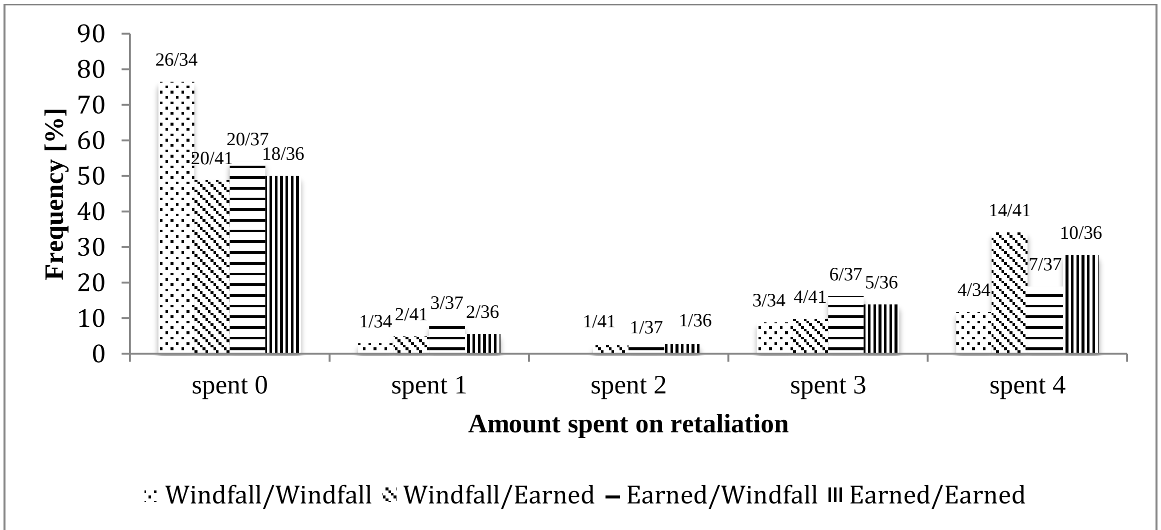


Table 1.3 Distribution of Final Payoffs

FM's Decision	SM's Decision	Total $ Surplus	Final $ Payoffs	Windfall/ Windfall (# pairs)	Windfall/ Earned (# pairs)	Earned/ Windfall (# pairs)	Earned/ Earned (# pairs)
Not Take $10	n/a	28	(10, 18)	3 (8.8%)	8 (19.5%)	9 (24.4%)	6 (16.7%)
Take $10	Spent $0	28	(20, 8)	23 (67.7%)	15 (36.6%)	17 (45.9%)	12 (33.3%)
Take $10	Spent $1	23	(16, 7)	1 (2.9%)	2 (4.9%)	2 (5.4%)	2 (5.6%)
Take $10	Spent $2	18	(12, 6)	0 (0%)	1 (2.4%)	1 (2.7%)	1 (2.8%)
Take $10	Spent $3	13	(8, 5)	3 8.8%)	3 (7.3%)	4 (10.8%)	5 (13.9%)
Take $10	Spent $4	8	(4, 4)	4 (11.8%)	12 (29.3%)	4 (10.8%)	10 (27.8%)

Hypothesis 1 states that there will be less retaliation in the Windfall/Earned treatment than in the Windfall/Windfall treatment. This comparison tests for the cost-of-reciprocity effect while controlling for the FM taking the windfall part of the SM's endowment, which is a weak violation of property rights. The hypothesis is based on a conjecture that if money used to pay for retaliation is earned as opposed to received from the experimenter as a windfall gain, one might perceive the retaliation to be more costly due to earned money being in a different mental account than windfall. Result 1 summarizes the finding.

Result 1: Increasing the cost of reciprocity results in ***more*** retaliation under the weak violation of property rights.

Support for Result 1: As can be seen from Table 1.2, the frequency and extent of retaliation is ***higher*** in the Windfall/Earned treatment than in the Windfall/Windfall treatment, i.e., subjects retaliated more with earned money than with windfall. The two-sided Fisher's exact test, reported in the first row of Panel A in Table 1.4, supports this observation by detecting that the frequency of retaliation is statistically significantly higher ($p = 0.018$) in the Windfall/Earned treatment than in the Windfall/Windfall treatment. The Mann-Whitney test reported in the same row detects a statistically significant difference between the two treatments in the extent of retaliation ($p = 0.012$), thus leading to the rejection of Hypothesis 1. The ordered probit regression reported in Panel A of Table1.5 supports this conclusion ($p = 0.008$).[10]

□

[10] Appendix A presents further analysis of the SM behaviour, namely a probit regression of a binary retaliation decision (Panel A) and an ordered probit regression of the level of retaliation conditional on retaliating (Panels B with marginal effects presented in Panel C).

Table 1.4 Tests for the Second Mover Behaviour

Panel A: Tests for the Second Mover Behaviour with Treatment Data

	The Mann-Whitney Test	Fisher's Exact Test
Windfall/Earned vs. Windfall/Windfall	- 2.511 (0.012)	(0.018)
Earned/Windfall vs. Windfall/Windfall	-1.795 (0.073)	(0.080)
Earned/Windfall vs. Windfall/Earned	0.926 (0.354)	(0.658)
Earned/Earned vs. Earned/Windfall	-0.607 (0.544)	(0.455)
Earned/Earned vs. Windfall/Earned	0.316 (0.752)	(0.548)
Earned/Earned vs. Windfall/Windfall	-0.244 (0.025)	(0.020)

Panel B: Tests for the Second Mover Behaviour with Earned Treatments Pooled Together [a]

	The Mann-Whitney Test	Fisher's Exact Test
Windfall/Windfall vs. Earned Treatments Pooled	-2.555 (0.011)	(0.010)

All tests are 2-sided. p-values in parentheses.

[a] Data from all Earned treatments were pooled together and compared to Windfall/Windfall.

In light of the original conjecture that earning money increases the cost of negative reciprocity, the fact that SMs retaliated more with their earned money is somewhat surprising. While the increase in retaliation does not necessarily indicate that the SMs considered the two parts of their endowment to be fungible, it suggests that negative reciprocity and property rights are inherently related. From the current design it appears that negative reciprocity is stronger as long as some part of the endowment was earned through exerting real effort.

Hypothesis 2 states that there will be more retaliation in the Earned/Windfall treatment than in the Windfall/Windfall treatment. I conjectured that while controlling for the SM retaliating with the windfall part of his endowment, taking the earned part of the SM's endowment would be considered a stronger violation of property rights than taking the windfall part. This would then lead to stronger retaliation by the SM.

Result 2: When the cost of reciprocity is low, the strong violation of property rights results in more retaliation than the weak violation.

Support for Result 2: The frequency of retaliation is weakly statistically significantly higher in the Earned/Windfall treatment than in the Windfall/Windfall treatment according to the Fisher's exact test reported in the second row of Panel A in Table 1.4 ($p = 0.080$). The corresponding Mann-Whitney test detects that the extent of retaliation is also higher ($p = 0.073$), thus providing further support for Hypothesis 2. The ordered probit regression also supports this result. The regression coefficient of 0.507 for the Earned/Windfall treatment (reported in the second row of Panel A in Table 1.5) represents a weakly statistically significant difference from the Windfall/Windfall treatment ($p = 0.089$).

□

Table 1.5 Ordered Probit Regression of Amount Spent on Retaliation by Second Movers

Panel A: Ordered Probit Regression

	pseudo-R^2	χ^2	Prob> χ^2
	0.024	8.06	0.045
Treatment	Coefficient	Z	P>z
Windfall/Earned	0.771	2.64	0.008
Earned/Windfall	0.507	1.70	0.089
Earned/Earned	0.678	2.27	0.023
Constant	0.687		

Panel B: Ordered Probit Amount Spent Marginal Effects

Treatment	Amount Spent=1	Amount Spent=2	Amount Spent=3	Amount Spent=4
Windfall/Earned	0.003 (0.453)	0.003 (0.215)	0.033 (0.006)	0.249 (0.009)
Earned/Windfall	0.003 (0.261)	0.002 (0.188)	0.023 (0.032)	0.159 (0.095)
Earned/Earned	0.003 (0.472)	0.002 (0.218)	0.029 (0.009)	0.218 (0.026)

Windfall/Windfall is the omitted control variable in both panels. Panel B presents the marginal effects of amount spent on retaliation across treatments. p-values in parentheses.

Combining Hypotheses 1 and 2 presents a prediction for the combined cost-of-reciprocity and property-rights-violation effect between the Windfall/Earned and Earned/Windfall treatments. In particular, if the part of the SM's endowment used for retaliation is earned, it becomes more costly to retaliate, which should in turn result in less retaliation. Similarly, if the FM takes the windfall part of the SM's endowment, representing a weak property right violation, it should also lead to less retaliation than if he took the earned part. According to Hypothesis 3 I should therefore observe more retaliation in the Earned/Windfall treatment than in the Windfall/Earned treatment.

Result 3: I find no evidence of the combined cost-of-reciprocity and property-rights-violation effect.

Support for Result 3: The Fisher's exact test reported in the third row of Panel A in Table 1.4 detects no difference in the frequency of SMs' retaliation between these two treatments (p = 0.658). The corresponding Mann-Whitney detects no difference in the extent of retaliation either (p = 0.354).

□

Result 3 thus suggests that earning a part of the endowment yields the same amount of negative reciprocity irrespectively of whether the earned part of the endowment was taken or used for retaliation. Result 3 combined with the fact that SMs retaliated more in both the Windfall/Earned and Earned/Windfall treatments than in the Windfall/Windfall treatment thus lead to a conclusion that property right entitlements might trigger a stronger negatively reciprocal response. However, subjects do not seem to distinguish between situations when they retaliate using earned money versus using windfall as long as a part of the initial endowment is earned.[11]

Hypothesis 4 concerns the cost-of-reciprocity effect again, but this time in a situation when the earned part of the SM's endowment was taken.

Result 4: Increasing the cost of reciprocity has no effect on retaliation under the strong violation of property rights.

[11] It is acknowledged that there may be a non-trivial interaction between the property right entitlements and the actual amount earned in the real effort task.

Support for Result 4: The Fisher's exact test reported in the fourth row of Panel A in Table 1.4 detects no difference in the frequency of SMs' retaliation between the Earned/Earned and Earned/Windfall treatments (p = 0.455). The corresponding Mann-Whitney detects no difference in the extent of retaliation (p = 0.544).

□

Finally, Hypothesis 5 states that there will be more retaliation in the Earned/Earned treatment than in the Windfall/Earned treatment as taking the earned part of the SM's endowment would be considered a stronger violation of property rights than taking a windfall gain. Controlling for the SM retaliating with the earned part of his endowment, this would in turn lead to more retaliation.

Result 5: When the cost of reciprocity is high, I find no evidence that the strong violation of property rights results in more retaliation than the weak violation.

Support for Result 5: The Fisher's exact test reported in the fifth row of Panel A in Table 1.4 detects no difference in the frequency of SMs' retaliation between the Earned/Earned and Windfall/Earned treatments (p = 0.548). The corresponding Mann-Whitney detects no difference in the extent of retaliation (p = 0.752).

□

I also report the comparison between the Windfall/Windfall and Earned/Earned treatments. Consistent with the previous results, the frequency and extent of retaliation in Earned/Earned is statistically significantly higher than in the Windfall/Windfall treatment (the

Fisher's exact test p = 0.020 and the Mann-Whitney test p = 0.025, both presented in the sixth row of Panel A in Table 1.4).

Since I find no statistical difference in SMs' behaviour in the three treatments in which at least a part of the initial endowment is earned (i.e., Windfall/Earned, Earned/Windfall, and Earned/Earned), I pool the data together and compare it to the Windfall/Windfall treatment. The frequency and extent of retaliation in the Pooled Earned treatments is statistically significantly higher than in the Windfall/Windfall treatment (the Fisher's exact test p = 0.010 and the Mann-Whitney test p = 0.011, both presented in Panel B of Table 1.4). The ordered probit regression reported in Panel A of Table 1.6 supports this result (p = 0.010).

Table 1.6 Ordered Probit Regression of Amount Spent on Retaliation by Second Movers with Earned Treatments Pooled Together

Panel A: Ordered Probit Regression Amount Spent

	pseudo-R^2	χ^2	Prob> χ^2
	0.021	7.03	0.008
Retaliation decision	Coefficient	z	P>z
Pooled Earned	0.655	2.59	0.010
Constant	0.688		

Panel B: Ordered Probit Amount Spent Marginal Effects

Treatment	Amount Spent=1	Amount Spent=2	Amount Spent=3	Amount Spent=4
Pooled Earned	0.014 (0.120)	0.006 (0.184)	0.051 (0.034)	0.170 (0.002)

Data from all Earned treatments were pooled together and compared to Windfall/Windfall.

The latter results thus provide further evidence that subjects retaliate more if they earn at least a part of their endowment but do not distinguish between retaliating with earned money or windfall (as long as a part of the initial endowment is earned). Furthermore, there is no change in negative reciprocity when only a part of the endowment is earned as opposed to when the entire endowment is earned, suggesting a binary rather than increasing relationship between the earned fraction of the endowment and negative reciprocity.

Contrary to my first conjecture that there will be more retaliation with windfall than with earned money I find that subjects actually retaliate more with their earned money in cases where at least a part of their endowment is earned. I find support for my second conjecture that if earned money is taken from subjects, they retaliate more because of the violation of their stronger property rights established by performing a real-effort task. My results thus points out that endowing subjects with windfall funds, absent of clearly established property rights, diminishes their negatively reciprocal responses.

To summarize, the current experiment provides evidence of the house money effect of the following kind: First, if at least a part of the endowment is earned, appropriating some of the endowment leads to more retaliation by the original owner, irrespectively of whether the appropriated part was earned or not. In such a case the earned and the windfall part of the endowment seem to be fungible. Second, if the entire endowment is earned, there is more retaliation than if the entire endowment is windfall, but earning both parts of the endowment (taken as well as used for retaliation) does not lead to more retaliation compared to earning only one part.

1.5 Demographics

1.5.1 The First Movers

As demographic data were elicited in the post-experiment questionnaire it might be interesting to test if any of these are statistically significantly influencing the independent variable TAKE. All variables that were potentially relevant were regressed and the results from an OLS regression are reported in Table 1.7 below.

Table 1.7 Demographics Analysis, the First Movers

	OLS Coefficient (p-value) Dependent variable TAKE
WINDFALL/EARNED	-0.091
	(0.334)
EARNED/WINDFALL	-0.154
	(0.103)
EARNED/EARNED	-0.026
	(0.784)
MALE	0.030
	(0.658)
AGE	-0.014**
	(0.036)
SIBLINGS	0.018
	(0.486)
ECON	0.044
	(0.518)
NON NZ	0.084
	(0.260)
RELATIVE INCOME	-0.057
	(0.130)
CITY SIZE	0.011
	(0.747)
LIVE WITH OTHERS	-0.005
	(0.816)
MONEY	0.000
	(0.266)
FINANCE STUDY	-0.001
	(0.429)
NO. OF PEOPLE KNOWN	0.026
	(0.190)
CONSTANT	1.286
	(0.001)

Run on StataSE 12.0. *, **, *** refer to statistical significance at the 10%, 5% and 1% levels, respectively.

The WINDFALL/WINDFALL treatment is omitted from the regression to use it as the comparison basis for treatment effects, and to avoid perfect correlation. MALE is a dummy variable that is 1 when subject reported their gender as male and 0 if female. Non NZ is a dummy variable that is 1 when subjects nominated a nationality of a country that was not New Zealand, out of this group approximately 63% of FMs and 61.7% of SMs nominated Asian nationality, and approximately 13% of FMs and 23.4% of SMs nominated European nationality. ECON is a dummy variable that is 1 when subjects nominated an area of study that was 'Economics' or 'Business Economics'. LIVE WITH OTHERS is the number of people that currently live in the subject's household, MONEY is the amount of dollars that subjects nominated as their monthly non-accommodation budget, FINANCE STUDY is the proportion that subjects nominated as the fraction of their monthly budget that they fund themselves. NO. OF PEOPLE KNOWN is the number subject knows in the session. RELATIVE INCOME is the income of subjects' parents in comparison to other families in New Zealand when subjects were 16 years of age. The larger the income of the family they state the larger the variable RELATIVE INCOME. Subjects had five possibilities to choose from, i.e. far below average, below average, average, above average, far above average. CITY SIZE is the size of the community where the subject has lived the most time of their life. Subjects could choose from the following four options: up to 2 000 inhabitants, 2 000 to 10 000 inhabitants, 10 000 to 100 000 inhabitants, more than 100 000 inhabitants. The more inhabitants they state the higher is the variable CITY SIZE. MONEY and FINANCE STUDY were included as the best measures available to control for income. The variable ECON was used to control for the behaviour of economics students.[12] They are likely to have learnt

[12] Marwell and Ames (1981)state in their article's title "Economists free-ride, does anyone else?"

about concepts such as rationality and backwards induction, and therefore tend to be more likely to play strategies closer to what standard game theory would predict. In this game it would therefore be predicted that economics students will TAKE $10 at all times. The coefficients for the treatments in Table 1.7 confirm that there is no statistically significant difference in TAKE between the treatments, meaning that the source of the part of the endowment that was taken away did not matter for the FMs. The only statistically significant variable at 10% level is variable AGE suggesting that older subjects were taking less often.

1.5.2 The Second Movers

Demographic data from the second movers were elicited in the post-experiment questionnaire. By running an OLS regression I can test if any of the demographic data affected the second movers decision on how many dollars to spend on retaliation. The OLS regression (presented in Table 1.8) was run on any variables that could have an impact on the second movers' behaviour.

In contrary to the FMs OLS results, in this case the treatments had and effect on the amount spent on retaliation by SMs, in particular WINDFALL/EARNED and EARNED/EARNED treatments. In these two treatments SMs were spending significantly more on retaliation than in the WINDFALL/WINDFALL treatment, which confirms the results from the Mann-Whitney test. Demographic data collected in the questionnaire did not have any effect on SMs' decisions on the amount spent on retaliation.

Table 1.8 Demographics Analysis, The Second Movers

	OLS Coefficient (p-value) Dependent variable Amount Spent
WINDFALL/EARNED	1.068**
	(0.012)
EARNED/WINDFALL	0.663
	(0.134)
EARNED/EARNED	0.948**
	(0.029)
MALE	0.368
	(0.230)
AGE	0.015
	(0.626)
SIBLINGS	0.161
	(0.246)
ECON	-0.037
	(0.905)
NON NZ	0.145
	(0.676)
RELATIVE INCOME	-0.163
	(0.345)
CITY SIZE	-0.172
	(0.288)
LIVE WITH OTHERS	0.158
	(0.145)
MONEY	0.000
	(0.336)
FINANCE STUDY	-0.001
	(0.867)
NO. OF PEOPLE KNOWN	-0.284
	(0.293)
CONSTANT	0.251
	(0.843)

Run on StataSE 12.0. *, **, *** refer to statistical significance at the 10%, 5% and 1% levels, respectively.

1.6 Conclusion

This paper experimentally investigates the extent of reciprocal reactions to unkind behaviour when (a part of) the initial endowment is earned by performing a real-effort task. I contribute to the literature on reciprocity by developing a design that allows to identify two reasons why the source of the decision maker's endowment might matter for his negatively reciprocal behaviour. Following from this notion there are two possible reasons for the house money effect: (1) the perceived cost of reciprocity and (2) the perceived strength of the property rights violation. To identify these issues I implement four treatments that differ in the source of endowment being taken by another player and/or used for retaliation. The treatments are nested in the Taking Game in which the FM has an opportunity to take $10 from the SM who can then retaliate. Based on mental accounting (Thaler, 1985) I conjecture that SMs place their earned money and windfall funds in two different mental accounts and as a result retaliate less using their earned money because it increases the costs of negative reciprocity. Similarly, if it is earned money that is taken from them, I conjecture that they retaliate more because of the violation of their property rights, which are stronger for the account that stores earned funds.

While I find support for the latter conjecture, I also find that subjects actually retaliate more with their earned money than with windfall as long as at least a part of their endowment is earned. However, as long as a part of their endowment is earned, subjects do not seem to distinguish between situations when they retaliate using earned money versus using windfall, suggesting that their main motivation is the violation of property rights established by performing the real-effort task.

From the perspective of the existing models of reciprocity (Cox, Friedman, and Gjerstad, 2007; Cox, Friedman, and Sadiraj, 2008; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006; Rabin, 1993) and unconditional distributional preferences (Bolton, 1991; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Fehr and Schmidt, 1999) the four treatments in the current experiment are isomorphic and the only difference between them is the theoretically irrelevant source of endowment. The distinctions implied by the design of the current experiment, however, are central to understanding reciprocal preferences.

Even though the above models of other-regarding preferences do not predict changes in behaviour, I find it useful to interpret the data against two of them in order to shed some light on the surprising results. In particular, in the context of Revealed Altruism theory (Cox, Friedman, and Sadiraj, 2008) the data suggest that if a part of the SM's endowment is earned, the FM's taking action decreases the SM's 'negative conditional altruism' (i.e., increases the SM's negative reciprocity) by more than in a situation when the SM's endowment is a windfall gain. In the context of inequality aversion (Bolton, 1991; Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999) an explanation used by a significant portion of experimental papers dealing with negative reciprocity and punishment, my findings suggest that the aversion to inequality increases when (a part of) the decision maker's endowment is earned by exerting real effort.

Based on the previous evidence that property right entitlements lead to more self-regarding behaviour, one could expect that negative reciprocity would be less pronounced when earned wealth is at stake, which is indeed what Bosman, Sutter, and van Winden (2005) find. The results from the current experiment stand in contrast with their Power-to-Take Game findings. In particular, Bosman, Sutter, and van Winden observe that the taking rate does not depend on effort and the responders destroy their endowment more often and in greater amount when effort is not involved – providing evidence for the overall house money effect. I find the incidence of taking to be lower only when comparing the Earned/Windfall treatment and the Windfall/Windfall treatment. Furthermore, in Bosman, Sutter, and van Winden's experiment negative reciprocity is higher in the no-effort treatment than in the real-effort treatment whereas I find negative reciprocity to be higher in the real-effort treatments.

While both the Taking Game and the Power-to-Take Game allow the FM to appropriate a part of the SM's endowment and the SM to retaliate, there exist significant differences in the structure of the two games and experimental procedures that could have contributed to different behaviour.[13] For example, different retaliation technologies combined with different action spaces could have resulted in varying strengths of incentives between the two experiments. Also the fact that in the current experiment only SMs exerted real effort by cutting posters, whereas in Bosman, Sutter, and van Winden's experiment both types of players had to earn their endowment through solving two-task optimization problems, could have contributed to respecting property rights by the FMs and more aggressive retaliation by the SMs. Finally, subject pool differences (Dutch versus New Zealand students) are always a possibility.

The observation that earned endowments sometimes lead and sometimes do not lead to stronger reciprocal responses, raises an interesting question of how (various levels of) property rights interact with reciprocity. The current paper as well as Cox and Hall (2010) point out that stronger property rights might increase the intensity of

[13] The differences between the Taking Game and the Power to Take Game are described in footnote 5

reciprocation, however, Cox, Servátka, and Vadovič (2014) do not find evidence of earned endowments affecting second movers reciprocal responses. It is therefore likely that the initial conditions (for example, modeled by the notion of status quo in the above mentioned Cox, Friedman, and Sadiraj, 2008) as well as the process how they originate are likely to have implications for reciprocal behaviour.

Appendix A. Further Analysis of the Second Mover Behaviour

Table A Additional Regressions Of The Second Mover Behaviour On Treatment Dummies

Panel A: Probit Regression of the Second Mover Binary Retaliation Decision

		pseudo-R^2	χ^2	Prob> χ^2
		0.038	7.59	0.055
Treatment	Coefficient	Marginal	z	P>z
Windfall/Earned	0.752	0.278	2.45	0.014
Earned/Windfall	0.620	0.228	1.97	0.048
Earned/Earned	0.722	0.266	2.29	0.022
Constant	-0.722	0.428	-3.05	0.002

Panel B: Ordered Probit Regression of the Level of Retaliation Conditional on Retaliating

	pseudo-R^2	χ^2	Prob> χ^2
	0.016	2.16	0.541
Treatment	Coefficient	z	P>z
Windfall/Earned	0.304	0.63	0.530
Earned/Windfall	-0.250	-0.51	0.608
Earned/Earned	0.071	0.15	0.884
Constant	-1.121		

Panel C: Ordered Probit Regression of the Level of Retaliation Marginal Effects Conditional on Retaliating

Treatment	Amount Spent=2	Amount Spent=3	Amount Spent=4
Windfall/Earned	-0.015	-0.046	0.118
	(0.547)	(0.557)	(0.525)
Earned/Windfall	0.012	0.033	-0.098
	(0.624)	(0.583)	(0.610)
Earned/Earned	-0.003	-0.010	0.027
	(0.884)	(0.886)	(0.884)

Appendix B. Subject Instructions

Windfall/Windfall Treatment

INSTRUCTIONS

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Anonymity

You have been randomly divided into two groups, called Group A and Group B. Each person in Group A will be randomly paired with a person in Group B. No one will learn the identity of the person (s)he is paired with.

Structure of the Experiment

Each Group A and Group B person will make only one decision in this experiment. The experiment is computerized. If you have any trouble entering your decisions on the computer, please raise your hand to alert the experimenter who will assist you.

Starting Balance

Each person in Group A will start with a starting balance of $10. Each person in Group B will start with a starting balance of $18.

Group A Decision Task

Each Group A person decides whether or not to take $10 from the starting balance of the paired Group B person.

If the Group A person decides **not to take $10** then the Group A person receives final earnings of $10 and the Group B person receives final earnings of $18.

If the Group A person decides **to take $10**, the paired Group B person then decides whether to decrease the Group A person's earnings. Group B person's decision is explained below.

Group B Decision Task

If the Group A person has decided to take $10 from the starting balance of the Group B person, the Group B person can decrease the Group A person's final earnings. Decreasing a Group A person's final earnings by $4 costs Group B person $1 which will be subtracted from his/her remaining $8. The following table shows five possible decisions that are available to a Group B person and the resulting final earnings for the pair.

	Group B person decides to decrease				
	Group A person's earnings by $0	**Group A person's earnings by $4**	**Group A person's earnings by $8**	**Group A person's earnings by $12**	**Group A person's earnings by $16**
Group A person's final earnings	**$20**	**$16**	**$12**	**$8**	**$4**
Group B person's final earnings	**$8**	**$7**	**$6**	**$5**	**$4**

Note that the decision by the Group B person will only be relevant if the Group A person chose to take $10 from Group B person.

Payment of Experiment Earnings

All participants are asked to sit patiently until the end of the experiment. Once all Group B persons have made their decisions, you will be presented with a summary screen of your earnings. Click OK after you have seen this screen, so other participants cannot see your decisions. You will then be prompted to complete a questionnaire. After the questionnaire, you will be asked one by one to approach the payment room at the back of the lab for the payment of your earnings. All the money will be paid to you in cash at the end of the experiment. Because your decision is private, we ask that you do not tell anyone your decision or your earnings either during or after the experiment. We also ask you to not gather near the lab after you receive your payment.

Are there any questions?

Windfall/Earned Treatment

TASK 1 INSTRUCTIONS

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

You have been randomly divided into two groups, called Group A and Group B. Each person in Group B will now have the opportunity to earn money.

Group B Task

In today's experiment each person in Group B will participate in a task, where (s)he will get a chance to earn $8. Each participant will be given 20 posters promoting NZEEL experiments. These posters need to be cut in a way that people passing by can take a stub with web page link where they can register for the experiments. The posters will then be placed in different parts of the university in order to recruit subjects for future experiments. Please cut the posters individually so that the stubs are neat. You will be paid only if you finish cutting all 20 posters that will be given to you.

Group A Has No Task

While Group B persons perform their task, we ask all Group A persons to wait patiently and quietly. Please do not use the computer in front of you as it is set up for the experiment.

Task 2 of the experiment will follow shortly.

TASK 2 INSTRUCTIONS

Anonymity

In Task 2 each person in Group A will be randomly paired with a person in Group B. No one will learn the identity of the person (s)he is paired with.

Structure of Task 2

Each Group A and Group B person will make only one decision in Task 2, which is the final part of the experiment. That is, after Task 2 there are no more tasks.

Task 2 is computerized. If you have any trouble entering your decisions on the computer, please raise your hand to alert the experimenter who will assist you.

Starting Balance

Each person in Group A as well as in Group B will start with a starting balance of $10. In addition to his/her starting balance each person in Group B has participated in a task, where (s)he earned $8.

Group A Decision Task

Each Group A person decides whether or not to take the $10 starting balance from the paired Group B person.

If the Group A person decides **not to take $10** then the Group A person receives final earnings of $10 and the Group B person receives final earnings of $18 ($10 starting balance and $8 from Task 1).

If the Group A person decides **to take $10**, the paired Group B person then decides whether to decrease the Group A person's earnings. Group B person's decision is explained below.

Group B Decision Task

If the Group A person has decided to take the $10 starting balance from Group B person, the Group B person can decrease the Group A person's final earnings using the money (s)he has earned in the Task 1. Decreasing a Group A person's final earnings by $4 costs Group B person $1 which will be subtracted from his/her Task 1 earnings of $8. The following table shows five possible decisions that are available to a Group B person and the resulting final earnings for the pair.

	Group B person decides to decrease				
	Group A person's earnings by $0	**Group A person's earnings by $4**	**Group A person's earnings by $8**	**Group A person's earnings by $12**	**Group A person's earnings by $16**
Group A person's final earnings	**$20**	**$16**	**$12**	**$8**	**$4**
Group B person's final earnings	**$8**	**$7**	**$6**	**$5**	**$4**

Note that the decision by the Group B person will only be relevant if the Group A person chose to take $10 from Group B person.

Payment of Experiment Earnings

All participants are asked to sit patiently until the end of the experiment. Once all Group B persons have made their decisions, you will be presented with a summary screen of your earnings. Click OK after you have seen this screen, so other participants cannot see your decisions. You will then be prompted to complete a questionnaire. After the questionnaire, you will be asked one by one to approach the payment room at the back of the lab for the payment of your earnings. All the money will be paid to you in cash at the end of the experiment. Because your decision is private, we ask that you do not tell anyone your decision or your earnings either during or after the experiment. We also ask you to not gather near the lab after you receive your payment.

Are there any questions?

Earned/Windfall Treatment

TASK 1 INSTRUCTIONS

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

You have been randomly divided into two groups, called Group A and Group B. Each person in Group B will now have the opportunity to earn money.

Group B Task

In today's experiment each person in Group B will participate in a task, where (s)he will get a chance to earn $10. Each participant will be given 20 posters promoting NZEEL experiments. These posters need to be cut in a way that people passing by can take a stub with web page link where they can register for the experiments. The posters will then be placed in different parts of the university in order to recruit subjects for future experiments. Please cut the posters individually so that the stubs are neat. You will be paid only if you finish cutting all 20 posters that will be given to you.

Group A Has No Task

While Group B persons perform their task, we ask all Group A persons to wait patiently and quietly. Please do not use the computer in front of you as it is set up for the experiment.

Task 2 of the experiment will follow shortly.

TASK 2 INSTRUCTIONS

Anonymity

In Task 2 each person in Group A will be randomly paired with a person in Group B. No one will learn the identity of the person (s)he is paired with.

Structure of Task 2

Each Group A and Group B person will make only one decision in Task 2, which is the final part of the experiment. That is, after Task 2 there are no more tasks.

Task 2 is computerized. If you have any trouble entering your decisions on the computer, please raise your hand to alert the experimenter who will assist you.

Starting Balance

Each person in Group A will start with a starting balance of $10. Each person in Group B will start with a starting balance of $8. In addition to his/her starting balance each person in Group B has participated in a task, where (s)he earned $10.

Group A Decision Task

Each Group A person decides whether or not to take the $10 which Group B person has earned in the previous task.

If the Group A person decides **not to take $10** then the Group A person receives final earnings of $10 and the Group B person receives final earnings of $18 ($8 starting balance and $10 from Task 1).

If the Group A person decides **to take $10**, the paired Group B person then decides whether to decrease the Group A person's earnings. Group B person's decision is explained below.

Group B Decision Task

If the Group A person has decided to take the earned $10 from Group B person, the Group B person can decrease the Group A person's final earnings using the money from the starting balance. Decreasing a Group A person's final earnings by $4 costs Group B person $1 which will be subtracted from his/her starting balance of $8. The following table shows five possible decisions that are available to a Group B person and the resulting final earnings for the pair.

	Group B person decides to decrease				
	Group A person's earnings by $0	**Group A person's earnings by $4**	**Group A person's earnings by $8**	**Group A person's earnings by $12**	**Group A person's earnings by $16**
Group A person's final earnings	**$20**	**$16**	**$12**	**$8**	**$4**
Group B person's final earnings	**$8**	**$7**	**$6**	**$5**	**$4**

Note that the decision by the Group B person will only be relevant if the Group A person chose to take $10 from Group B person.

Payment of Experiment Earnings

All participants are asked to sit patiently until the end of the experiment. Once all Group B persons have made their decisions, you will be presented with a summary screen of your earnings. Click OK after you have seen this screen, so other participants cannot see your decisions. You will then be prompted to complete a questionnaire. After the questionnaire, you will be asked one by one to approach the payment room at the back of the lab for the payment of your earnings. All the money will be paid to you in cash at the end of the experiment. Because your decision is private, we ask that you do not tell anyone your decision or your earnings either during or after the experiment. We also ask you to not gather near the lab after you receive your payment.

Are there any questions?

Earned/Earned Treatment

TASK 1 INSTRUCTIONS

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

You have been randomly divided into two groups, called Group A and Group B. Each person in Group B will now have the opportunity to earn money.

Group B Task

In today's experiment each person in Group B will participate in a task, where (s)he will get a chance to earn his/her starting balance of $18. Each participant will be given 20 posters promoting NZEEL experiments. These posters need to be cut in a way that people passing by can take a stub with web page link where they can register for the experiments. The posters will then be placed in different parts of the university in order to recruit subjects for future experiments. Please cut the posters individually so that the stubs are neat. You will be paid only if you finish cutting all 20 posters that will be given to you.

Group A Has No Task

While Group B persons perform their task, we ask all Group A persons to wait patiently and quietly. Please do not use the computer in front of you as it is set up for the experiment.

Task 2 of the experiment will follow shortly.

TASK 2 INSTRUCTIONS

Anonymity

In Task 2 each person in Group A will be randomly paired with a person in Group B. No one will learn the identity of the person (s)he is paired with.

Structure of Task 2

Each Group A and Group B person will make only one decision in Task 2, which is the final part of the experiment. That is, after Task 2 there are no more tasks.

Task 2 is computerized. If you have any trouble entering your decisions on the computer, please raise your hand to alert the experimenter who will assist you.

Starting Balance

Each person in Group A will start with a starting balance of $10. Each person in Group B will start with a starting balance of $18, which (s)he earned in Task 1.

Group A Decision Task

Each Group A person decides whether or not to take $10 which Group B person has earned in Task 1.

If the Group A person decides not to take $10 then the Group A person receives final earnings of $10 and the Group B person receives final earnings of $18 which (s)he earned in Task 1.

If the Group A person decides to take $10, the paired Group B person then decides whether to decrease the Group A person's earnings. Group B person's decision is explained below.

Group B Decision Task

If the Group A person has decided to take $10 from Group B person's earned starting balance, the Group B person can decrease the Group A person's final earnings using the remaining money (s)he has earned in the Task 1. Decreasing a Group A person's final earnings by $4 costs Group B person $1 which will be subtracted from his/her remaining Task 1 earnings of $8. The following table shows five possible decisions that are available to a Group B person and the resulting final earnings for the pair.

	Group B person decides to decrease				
	Group A person's earnings by $0	Group A person's earnings by $4	Group A person's earnings by $8	Group A person's earnings by $12	Group A person's earnings by $16
Group A person's final earnings	$20	$16	$12	$8	$4
Group B person's final earnings	$8	$7	$6	$5	$4

Note that the decision by the Group B person will only be relevant if the Group A person chose to take $10 from Group B person.

Payment of Experiment Earnings

All participants are asked to sit patiently until the end of the experiment. Once all Group B persons have made their decisions, you will be presented with a summary screen of your earnings. Click OK after you have seen this screen, so other participants cannot see your decisions. You will then be prompted to complete a questionnaire. After the questionnaire, you will be asked one by one to approach the payment room at the back of the lab for the payment of your earnings. All the money will be paid to you in cash at the end of the experiment. Because your decision is private, we ask that you do not tell anyone your decision or your earnings either during or after the experiment. We also ask you to not gather near the lab after you receive your payment.

Are there any questions?

Appendix C. Control Questions

1. What will be the Group A person's final earnings if (s)he does not take $10?

2. What will be the Group B person's final earnings if the Group A person does not take $10?

3. What will be the Group B person's final earnings if the Group A person takes $10 and the Group B person does not decrease the Group A person's earnings?

4. What will be the Group A person's final earnings if (s)he takes $10 and the Group B person does not decrease the Group A person's earnings?

5. What will be the Group B person's final earnings if the Group A person takes $10 and the Group B person decreases the Group A person's earnings by $16?

6. What will be the Group A person's final earnings if (s)he takes $10 and the Group B person spends $3 of his/her own earnings on such decrease?

7. What is the cost to the Group B person of decreasing the paired Group A person's earnings by $8?

Appendix D. The Questionnaire

Earned/Earned Treatment

The following questions were asked either Group A or Group B persons and their decision determined, which of these they were asked:

Group A person:

The money that you took from the Group B person was earned/not earned by them.

Why did you decide to take $10 from Group B person?

The money that you could have taken, but didn't, from the Group B person was earned/not earned by them.

Why did you decide not to take $10 from the Group B person?

Group B person:

The money that the Group A person took from you was: earned/not earned by you.

Why do you think the Group A person took $10 from you?

The money that the Group A person could have, but didn't take from you was earned/not earned by you.

Why do you think the Group A person didn't take $10 from you?

Questions asked to all the participants:

Please state your gender.

How old are you?

What is your nationality?

How many siblings do you have?

If you are a student, what is your subject?

When you were 16 years of age, what was the income of your parents in comparison to other families in New Zealand?

How large was the community where you have lived the most time of your life?

How many people live in your household (including yourself)?

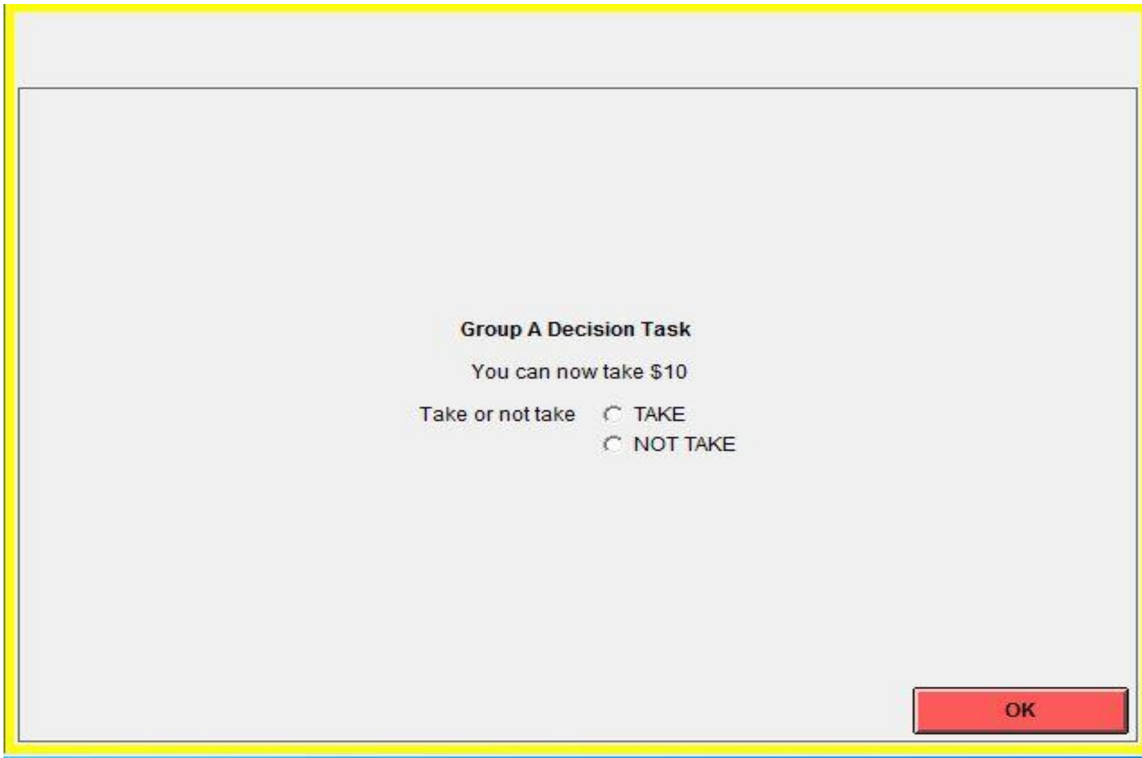
How large is your monthly budget (without expenses for accommodation)?

What share of your monthly expenses do you finance yourself?

Can you state the percentage we can rely on the data you provided?

Appendix E. Screenshots from the software

Earned/Windfall Treatment



Group B Decision Task

Decreasing the Group A person's earnings by $4 will cost you $1.

This amount will be subtracted from starting balance of $8.

How much money do you want to subtract from your starting balance in order to decrease your paired Group A person's earnings?

- 0
- 1
- 2
- 3
- 4

Note that this decision will only be relevant if the Group A person chose to take $10 from you.

OK

Your paired person took your earned money.

Your starting balance was	8
You have earned	10
Your paired person took	10
Your earnings before decreasing the paired Group A person's earnings	8
The amount you decreased your own earnings	4
Your final earnings are	4

OK

Appendix F. Human Ethics Committee Approval

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Ref: HEC 2013/46

28 May 2013

Katarina Dankova
Department of Economics & Finance
UNIVERSITY OF CANTERBURY

Dear Katarina

The Human Ethics Committee advises that your research proposal “House money effect and negative reciprocity” has been considered and approved.

Please note that this approval is subject to the incorporation of the amendments you have provided in your email of 20 May 2013.

Best wishes for your project.

Yours sincerely

Lindsey MacDonald
Chair
University of Canterbury Human Ethics Committee

Chapter 2

2 Job Assignment, Transparency and Social Comparisons

2.1 Introduction

Consider an organization with two agents and two job slots; one low-productivity job slot (henceforth low slot) and one high-productivity job slot (henceforth high slot). The principal has to assign one of the two agents to the low slot and the other to the high slot. Assuming the principal equates the marginal product of labour to wages, and that interpersonal concerns are absent, the principal would provide a higher wage to the agent in the high slot. Since agents' perception of fairness may depend on the wage paid to the other agent (Frank, 1984) and on the intentions of the principal (Rabin, 1993), the agent in the low slot might feel that the wage disparity is unfair, perhaps because the principal favours his co-worker for reasons unrelated with the job, and he may withhold effort in response, thus reducing total output.

In this chapter I explore whether and how agents' concerns for equity depend on fairness of the assignment process to a slot in organizational setting. In particular, I am interested in (i) how agents respond to inequity in job assignments and (ii) whether principals selectively engage in wage compression to account for agents' responses to inequitable outcomes. These questions have implications for labour market, in particular a firm's wage policy. If an unfair assignment, e.g. assigning a more qualified agent to a lower paid job, elicits more wage rejections due to social comparisons, it can have detrimental effects on the performance of the firm. My main contribution lies in examining the impact of job assignment -- an important feature of labour relationships --

in the presence of social comparisons (Bewley, 1999; Charness and Kuhn, 2007). Specifically, I compare agents' responses to random inequitable assignments, versus agents' responses to deliberate inequitable assignments by the principal.

Laboratory experiments are suitable for testing employment theory because they allow for the extensive control of decision environments, which is often difficult in naturally occurring settings. Wage theories have important implications for the functioning of labour markets, and can explain phenomena such as rigid wages and involuntary unemployment. Workers in firms are confronted with a mix of different incentives, which makes interpretation of their decisions difficult. An observed variation in wages may not reflect generosity but may be due to firm size, self-selection of workers, or simply productivity differences. In the laboratory, the experimenter can rule out confounding effects such as multiple incentives, selection, productivity differences, or repeated interactions. The experimenter also controls for payoffs, the order in which the different parties can act, and the information they receive when making their choice. This control allows for the testing of precise predictions derived from game-theoretic models (Falk and Heckman, 2009).

In the current experiment with one principal and two agents there are two different slots, which differ in the pie size. Each agent has to be assigned to one of the slots, in one treatment this is done randomly, in other two treatments it is done by the principal. Only after this assignment the principal decides on the wage for each agent. Prior to the assignment to slots the agents participate in a general knowledge quiz. The purpose of the quiz is to generate information, which can be used for comparisons of the two agents. I vary the process of assignment to a slot as well as the information the agents receive about the principal knowing or not knowing the relative quiz performance. Through varying agents' information I gain insights into the nature of fairness considerations underlying behaviour. Based on the findings from Charness (2004) I conjecture that if the agents care about the principal's intentions, they will reject wages more often in the case where the principal is responsible for the assignment to a slot, as opposed to a random assignment. In the case when the principal knows the relative quiz performance and assigns the agents to the slots, it is likely that he will assign the higher performer to the high slot and the lower performer to the low slot. This would not enable to study agents' reactions to unfair assignment, which might lead to stronger reciprocal responses such as rejecting wages more often than in the case of fair assignment. Therefore I add a scenario, in which the principal does not know the relative quiz performance and the agents do not receive the information about the principal not knowing the relative quiz performance. The agents, however, do receive information about the relative quiz performance. The principal thus decides on the assignment quasi randomly, as he has to assign the agents to the slots, but he does not know the relative quiz performance (or any other distinguishing information) to help him make his decision. This results in more agents with a higher relative quiz performance being assigned to low slots. Based on the literature on real effort and relative performance triggering strong entitlements (e.g. Bosman, Sutter, and van Winden, 2005; Cherry, Frykblom, and Shogren, 2002; Oxoby and Spraggon, 2008) I posit that such a situation is likely to be perceived as unfair by agents and might trigger stronger reciprocal reactions, therefore allowing agent's reactions to principal's intentions to be studied.

The current experiment is nested in a three-player ultimatum game with one principal and two agents. The experiment consists of three treatments; in all of them the agents take part in a general knowledge quiz and are informed about the relative quiz performance. In the first treatment (the Random Assignment treatment) the assignment to the low and high slots is random. In both the second (the Principal Assignment treatment) and the third (the Quasi-Random Assignment treatment) treatments the principal decides on the assignment and he knows the relative quiz performance of the agents in the Principal Assignment treatment but has no such information in the Quasi-Random Assignment treatment. In both of these treatments the agents know that the principal makes the assignment. In the Principal Assignment treatment they are explicitly told that the principal knows the relative quiz performance, but in the Quasi-Random Assignment treatment this information is omitted. This design results in a quasi-random assignment. From the agents' point of views, I test whether the agents compare their wages with the wages offered to the paired agents. From the principals' point of view, I test whether the principal reacts to the rankings from the quiz and whether he adjusts wages based on his beliefs about agents' perception of fairness. In other words, I test if there is "wage compression" in the Principal Assignment treatment relative to the Random Assignment treatment.

2.2 Literature Review

There is a great deal of literature on social comparisons, as they apply to the labour market, inspired by Bewley (1999) and surveyed by Charness and Kuhn (2007). Bewley (1999) points out that loyalty is highly valued among firms because hiring new workers is time-consuming and costly. To encourage loyalty firms often do not lower wages or they avoid layoffs during the economic downturns, resulting in downward wage-rigidity. Besides payoff considerations loyalty is often enhanced when workers are being treated fairly compared to their co-workers. Firms have to pay attention to the fact that individuals tend to evaluate outcomes in comparison to a reference point they consider fair (Adams, 1965; Deutsch, 1975; Leventhal, 1976a). The formation of a reference point depends on a measure of similarity between persons, i.e. the workers tend to compare their wages, their workspace and other benefits with their co-workers (Festinger, 1954; Jones, 1985). If a worker differs significantly from his co-worker, he is less likely to compare himself to that person. In order to avoid the potentially detrimental effects of social comparisons, in particular wage comparisons, organizations often engage in wage secrecy (Lawler, 1990) and wage compression to reduce differences in wages between workers (Akerlof and Yellen, 1990).

Economic experiments provide direct evidence that effort choices in a gift exchange game are affected by wage comparison (Güth et al., 2001; Requate, Waichman, and Siang, 2011). Clark, Masclet, and Villeval (2010) combine experimental evidence from a gift-exchange game and survey data from International Social Survey Program to test if individual effort depends on both one’s own income and an individual’s position in relative income distribution. The survey data analysis helps to check the external validity of experimental evidence. Their benchmark treatment consists of a gift-exchange game, in which a firm offers a wage to the worker and the worker then decides whether to accept or reject the wage. If the contract is rejected, both the firm and the worker receive nothing. Upon acceptance, the worker chooses his effort level, which is costly. The higher the effort level, the higher the profits of the firm and at the same time the higher the costs to the worker. In the information treatment, information about wage offers to the other workers is added to the first stage of the game, i.e. workers can compare their own wage to wages offered to workers in different firms (participating in the same experimental session) before rejecting or accepting the wage offer and choosing effort. They find that effort depends on one's own wages as well as on the wage of others. Another finding is that; (a) an individual's rank in the income distribution determines effort more strongly than others' average income, suggesting that comparisons are more ordinal than cardinal and (b) those who received a higher income or a higher income rank in the past exert less effort in the present. In my experiment, besides focusing on social comparisons between workers, I also test for the effect of intentions of the employer on workers' decisions.

Some studies find no evidence of social comparisons in the gift-exchange game. For example, Charness and Kuhn (2007) in experimentally test whether co-workers wages influence workers effort and find that while workers' effort choices are highly sensitive to their own wages, effort is not affected by co-workers' wages.

Wage comparisons have also been studied in the ultimatum game. A typical finding in ultimatum games is that agents reject a substantial fraction of low offers (under 20%) and proposers offer 40% of the pie on average (Camerer, 2003).[14] These findings are robust to stake size (Slonim and Roth, 1998) and to culture differences (Costa-Gomes and Zauner, 2001; Henrich, 2000; Henrich et al., 2001). When pie sizes are unknown to agents, principals make lower offers and agents accept lower offers. For example, Güth, Huck, and Ockenfels (1996) find that when pies sizes differ and are unknown to the agents, principals with large pies were pretending that the pie was small and offer a fair

[14] Hereafter „a proposer" is called a principal in the text to stress the focus of the paper, which is on the labour relations.

division in view of a small pie. Kagel, Kim, and Moser (1996) find that agents accept less when uneven offers are generated by a random device than by a principal. Charness et al. (2013) in their experiment on wage delegation find that workers are concerned both with their own salaries and their relative wages with respect to their co-workers and make lower effort choices when they cannot choose their own wage while their co-workers can.

In a three-player ultimatum game agents' rejection rate increases with the difference between the offer to the agent and the offer to the other agent. Knez and Camerer (1995) introduced a three-player ultimatum game with outside options. They find that agents reject offers more frequently if they are offered less than the other agent but principals do not seem to take this into consideration and do not adjust the offers. Ho and Su (2009) analyse two independent ultimatum games played sequentially by a principal and two agents to test if agents look at the other agents as a reference to evaluate their wage. They introduce a term "peer induced fairness", which means that the second agent is averse to receiving less than the first agent. The principal plays an ultimatum game with the first agent, and then the same principal plays an ultimatum game with the second agent. The second agent does not perfectly observe the offer the first agent receives, he only obtains an informative but imperfect public signal of the first offer, and can use this signal to infer the first agent's payoff. The signal was constructed by drawing a number from a discrete uniform distribution over the set {-20, -10, 0, 10, 20} and added to the first offer. Consequently, given a signal, second agents can infer what the first agent is likely to receive. The second agent is asked to make a guess of what the first offer is and is rewarded a sum of ten points for making a correct guess. This allows the second agent to have peer-induced fairness concerns and also allows for analysing the equilibrium of the game under imperfect information. Without peer-induced fairness, the second agent's acceptance decision and the principal's offer in the second game should not be influenced by the signal. The model predicts that the second agent's likelihood of accepting an offer decreases as the signal increases, suggesting that an identical offer can become less attractive as the second agent's expectations of the first agent's payoff increases. In addition, the principal's offer to the second agent is contingent on the signal. The higher the signal, the more the principal will offer the second agent. Ho and Su (2009) find support for their predictions, i.e. they find that the second agent's rate of rejection increases with the difference between the second offer and his expectations of the first agent's payoff. The second agent rejects the offer more frequently as the received signal increases. They also find that the principal aligns the second offer close to the expectation of the first agent's offer in order to avoid rejection by the second agent. The principal's offer is strategic in that he exploits the second agent when the signal is low and offers more when the signal is high. These results strongly suggest the existence of peer-induced fairness.

Apart from social comparisons, the perception of fairness is has been shown to depend also on intentions. Rabin (1993) argues that people differentiate between an intentionally mean act, which they may punish, and an unintentionally mean act, which they may tolerate. This theoretically demonstrates that intentions might matter. An intentional act might trigger social comparisons among agents, they might feel that they are being treated unfairly and this might lead to lower output of the firm. Principals thus need to take into account that agents are sensitive to intentions. There exists experimental evidence showing that agents attribute a different degree of intentionality to principal's decisions (Charness, 2004; Falk, Fehr, and Fischbacher, 2008).

Falk, Fehr, and Fischbacher (2008) use a moonlighting game to provide experimental evidence for the behavioural relevance of fairness intentions. The moonlighting game is a two-player (A and B) game that consists of two stages. Both players are endowed with 12 points. In the first stage, player A chooses an action $a \in \{-6, -5, ..., 5, 6\}$. If A chooses a >=0, he gives player B *a* tokens and the experimenter triples *a* so that B receives *3a*. If A chooses $a < 0$, he takes |a| tokens away from B, A gain $|a|$ and player B loses $|a|$. In the second stage, after player B observes *a*, he can choose an action $b \in \{-6, -5, . . . , 17, 18\}$, where $b >= 0$ is a reward and $b < 0$ is a sanction. A reward transfers *b* points from B to A. A sanction costs B exactly $|b|$ but reduces A's income by $3|b|$. Since As can give and take while Bs can reward or sanction, this game allows for both positively and negatively reciprocal behaviour. The experiment was run using the strategy method, i.e. player B had to give a response for each feasible action of player A, before B was informed about A's actual choice. The treatments differ in whether A's choice is under his full control. A determines *a* in the Intention treatment, if *a* is high, his action signals intentional kindness and if *a* is low, it signals intentional unkindness. In contrast, a random device (casting two dice) determines A's move in the No-intention treatment. Consequently, A has no control over his action. His action therefore signals neither good nor bad intentions. The main result of the experiment indicates that the attribution of fairness intentions is important in both the cases of negatively and positively reciprocal behaviour. When the experimental design rules out the attribution of fairness intentions, reciprocal responses are substantially weaker.

When wages differ according to performance, the principal does not decide only on the wage level but also on the relative wages and whether or not the high performer is rewarded by a higher wage compared to his low performance co-worker. For example, Abeler et al. (2010) observe in a gift exchange game with a principal and two agents that agents who are paid equal wages exert significantly lower efforts than the agents who are paid individually. In Charness (2004), the wage in the gift-exchange game is either chosen by the principal or by external process (either a draw from a bingo cage or an assignment made by the experimenter). The game was played for ten rounds with random matching of the principal and agent to avoid reputation concerns. After the principal assigned the wage, the agents were asked to record their effort choices. Charness (2004) finds that the provided effort depends on both the wages and the mechanism for its determination. In the current experiment, however, the wage is always chosen by the principal and it is the process of assignment to a slot that varies across treatments. The principal decides on the wage for each agent. Each agent is informed which slot he has been assigned to and is also informed about the other agent's wage. Agents can thus react to the assignment to the slot and to the other agent's wage.

Furthermore, if it is possible for agents to attribute a different degree of intentionality to principals' wage choices, wages are significantly more likely to be rejected when principals intentionally offer unequal money splits compared to when comparable wages are clearly unintentional. For example, Kagel, Kim, and Moser (1996) try to gain insight into the nature of fairness considerations underlying behaviour, by varying players' information and payoffs. In their ultimatum game, players bargain over

100 chips, where the chips have different monetary payoffs to the two players ($0.10 or $0.30 per chip). In addition, players have different information regarding chip payoffs: In some treatments only one player knows both payoffs and in others both players know both payoffs. In all cases, players know their own payoffs and it is "common knowledge" whether players know each other's payoffs. Offers are significantly more likely to be rejected when first-movers intentionally offer unequal money splits compared to when comparable offers are clearly unintentional (due to lack of knowledge). When both players are fully informed and first-movers have higher exchange rates, conflicting fairness norms developed, resulting in unusually high rejection rates.

Gächter and Thöni (2010) present three experiments, all of them using a three-person gift-exchange game, investigating the impact of wage comparisons for agent productivity. In the first experiment, the game is played in direct response mode, repeated eight times with randomly re-matched groups of three players in each round.[15] They find that agents who face disadvantageous wage discrimination significantly reduce their effort relative to a situation with equal wages. In the second experiment the strategy method is used to allow focus on individual differences in wage comparisons. This has the advantage of increasing observations and allows for testing for individual heterogeneity. The result of this experiment also shows that there are significant wage comparison effects in the case of disadvantageous wage discrimination. Consistent with the fair wage-effort hypothesis agents tend to reduce their effort when they are paid less than the other agents but no equivalent effect is observed for overpayment. These average effects mask a surprisingly large variety of different patterns of wage comparison effects.

[15] In a direct-response method, a second mover learns the action of the first mover and then chooses a response whereas in a strategy method, a second mover makes conditional decisions for each possible information set.

Some agents seem to reward the principal for a generous wage for the other agent, while others do the opposite. In their third experiment, which is the most relevant for the topic of this chapter, Gächter and Thöni (2010) test whether wage comparison effects are due to intentional wage discrimination or due to wage differences. The game is played using the strategy method and principals are not responsible for discriminatory wages, rather the wages are chosen randomly. They find that it is not the wage consequence that triggers the wage comparison effect but the procedure used to arrive at that wage. If intentions do not matter, the offered wages should lead to the same level of effort, regardless of the intentional/unintentional process behind it, which is not what is observed in Gächter and Thöni (2010) and in Falk, Fehr, and Fischbacher (2008).

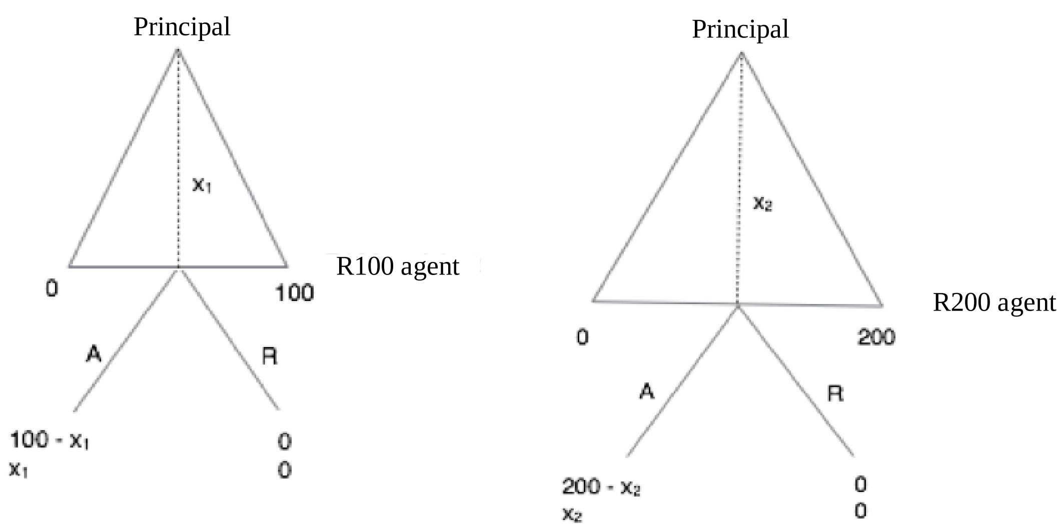
While I combine the findings on intentions and social comparisons in the principal-agent relationship, I distinguish between two different levels of job slots, i.e. one that could possibly pay more and one that could pay less. I conjecture that an agent's reaction to a specific slot assignment depends on the procedure used to arrive at that outcome. An ex-post assignment to a low slot will not be perceived as unfair if it arises from an ex-ante unbiased procedure, for example random assignment with equal probabilities, but will be perceived as unfair if agents believe that the procedure is biased; in particular, if the principal assigns the agent with a higher relative quiz performance to a low slot. To the best of my knowledge these questions have not yet been answered in the previous literature.

I address these questions using a three-player ultimatum game, in which a principal and each agent divide a fixed pie. The principal moves first and offers a division of the pie to one agent, and a division of a separate pie to the other agent (wage). The agents can accept or reject the wage. If the agent accepts, the pie is distributed according to the proposal. If an agent rejects, both the agent and the principal earn nothing.

2.3 Experimental Design And Procedures

This experimental design is nested in a three-player ultimatum game, presented in Figure 2.1, in which one principal (called proposer in the experiment in order to avoid framing effects) is paired with two ex-ante identical agents: agent A and agent B. The computer randomly determines who is the principal, agent A and agent B. Each participant has a 1/3 chance of becoming the principal, agent A, or agent B. All interactions are anonymous. In the first part of the experiment agents are asked to complete a general knowledge quiz. Each agent is asked to answer the same set of 20 questions in the same order and each question has one correct answer. While agents complete the quiz, all principals are asked to wait patiently and quietly.

Figure 2.1 Game Tree



Within each group, the two agents are ranked based on their quiz scores. If both agents have the same score in the quiz, the agent who completes the quiz the fastest is ranked higher. After the completion of the quiz the agents are assigned to the low slot and the high slot (in the instructions represented by the R100 and R200 roles, respectively, to avoid framing effects). After the assignment, the agents are labelled the R100 agent if assigned to the low slot and the R200 agent if assigned to the high slot. The assignment varies depending on the experimental treatment and is always common knowledge. There are three treatments (see Table 2.1). In the Principal Assignment treatment the principal and both agents are informed about the relative quiz performance. The principal decides on the assignment to a slot and on the wages based on the relative quiz performance. In the Random Assignment treatment, the principal and both agents are also informed about the relative quiz performance but the assignment to a slot is random. The principal decides only on the wages based on the relative quiz performance. In the Quasi-Random Assignment treatment only the agents are informed about the relative quiz performance. The principal decides on the assignment to a slot and on the wages without being informed about the relative quiz performance. I posit that the principal will assign the agent who had a higher relative quiz performance to the high slot. In the Quasi-Random Assignment treatment the principal has to make the assignment without knowing the relative quiz performance. The fact that the principal does not know the relative performance is withheld from the agents in the Quasi-Random Assignment treatment in order to control for the change in the agents responses (to unfair assignment) compared to the Principal Assignment treatment.

The principal and the R200 agent receive a 200-franc pie to be divided between them.[16] Separately, the principal and the R100 agent receive a 100-franc pie to be divided between them. The procedure for dividing each pie between the principal and each agent is as follows. The principal chooses how many francs out of 200 to offer to the R200 agent and how many francs out of 100 to offer to the R100 agent.

Agents' responses are elicited using strategy method, hence each agent does not observe the wage that the principal offered him; however, he observes the wage that the principal offered the paired agent in order to generate social comparisons.[17] After observing the wage that the principal offered to the paired agent, each agent states his minimum acceptable wage (hereafter minimum acceptable offer, MAO). Each agent chooses his MAO before he comes to know his actual wage. If the wage offered by the principal turns out to be greater or equal to the MAO, the wage is accepted. If the wage is accepted, the agent receives the number of francs stated in the wage, whereas the principal keeps the remainder. However, if the wage is less than MAO then the wage is rejected, the pie disappears and both the principal and the agent receive nothing.

Each observation consists of two wages offered by the principal (to the R100 agent and to the R200 agent) and a MAO stated by each of the two agents. Each decision is statistically independent because an agent's decision on the MAO does not affect the paired agent's decision or monetary payoff. This allows me to test whether social comparisons occur between the two paired agents.

[16] 1 franc = 0.1 NZD

[17] Armantier (2006) did not find differences in MAOs in an ultimatum game using strategy method versus direct response method (listed in Brandts and Charness (2011)).

Table 2.1 Overview Of The Treatment Differences

	Random Assignment	Principal Assignment	Quasi-Random Assignment
Assignment to the high and low slots	Random	Principal	Principal
The principal knows about the relative quiz performance	Yes	Yes	No

I conjecture that the sensitivity to the paired agent's wage will be greater if the assignment to the slots is done by the principal as opposed to random assignment. In the Principal Assignment treatment the principal is responsible for the wage and also for the assignment of the agents to the slots. The agents might thus attribute intentions to the principal's decisions, compare themselves with their paired agent and state MAOs that are closer to the wage of the paired agent than in the Random Assignment treatment. I expect the wage the principal offers to the R100 agent to be lower than the wage offered to the R200 agent, because of differing pie sizes (100 francs and 200 francs). I also expect larger differences in wages to the R100 and R200 agents in the Random Assignment treatment (offering more to the R200 agent and offering less to the R100 agent) whereas in the Principal Assignment treatment I expect the principal to offer the wage to both agents. In the Random Assignment treatment the principal is not responsible for assigning the agents to the slots, the only thing he can do is to decide on wage given the fact that the agents have already been assigned to their slots. I would thus expect the agents to be more tolerant to the wage offered by the principal, thus the agents' MAO and the paired agents' wage differentials will be higher than in the Principal Assignment treatment. The principal anticipating this will offer different wages to the R100 agent and

R200 agent. In the Principal Assignment treatment, however, the principal has to assign the agents to the slots as well as choose the wage. The size of the pie, which is available for each agent and the principal, thus depends on the principal's decision, to which the agents might attribute intentions. I expect the principal to expect the R100 agent to state his MAO closer to the R200 agent's wage than in the Random Assignment treatment resulting in offering similar wages to both agents, i.e. compressing the wages. I test the following hypotheses:

Hypothesis 1: The R100 agents' MAO and the R200 agents' wage differentials are smaller in the Principal Assignment treatment than in the Random Assignment treatment.

Hypothesis 2: The R100 agents' MAO and the R200 agents' wage differentials are smaller in the Quasi-Random Assignment treatment than in the Random Assignment treatment.

Hypothesis 3: The R200 agents' MAO and the R100 agents' wage differentials are smaller in the Principal Assignment treatment than in the Random Assignment treatment.

Hypothesis 4: The R200 agents' MAO and the R100 agents' differentials are smaller in the Quasi-Random Assignment treatment than in the Random Assignment treatment.

These hypotheses test whether the agents' MAO approximates to the paired agents' wage, suggesting that the paired agents' wage serves as a reference point. In the

Random Assignment treatment, the assignment to the slots is random; the principal only decides on the wages. Based on previous literature I posit that the MAOs are more likely to be closer to the paired agents' wages in the Principal Assignment treatment than in the Random Assignment treatment (Charness, 2004; Kagel, Kim, and Moser, 1996). I thus expect the agent's MAO and the paired agent's wage differentials to be smaller in the Principal Assignment treatment. In the Quasi-Random Assignment treatment there is a higher possibility of agents whose relative quiz performance was higher, to be assigned to a low slot, which might trigger stronger responses in the form of smaller differentials between the agent's MAO and the paired agent's wage than in the Random Assignment treatment. Since agents in the Quasi-Random Assignment treatment are not informed about the principal not knowing the relative quiz performance, they might perceive that the assignment to the low and high slots as well as the wages are intentionally unfair. Thus, agents might not be as "forgiving" as in the Random Assignment treatment and might expect higher wages and state higher MAOs, which in turn leads to more rejections – similarly as in the Principal Assignment treatment.

Hypothesis 5: The R100 agents' wage and the R200 agents' wage differentials will be smaller in the Principal Assignment treatment than in the Random Assignment treatment.

If agents' behaviour depends on the assignment procedure and agents are sensitive to the other agent's wage, the principal reacts by compressing wages in the Principal Assignment treatment. The assignment to slots in the Principal Assignment treatment is done by the principal, who knows the relative quiz performance and the agents are aware of it. If the wages differ significantly, the agents might feel this is unfair and reject the wages more often. In order to avoid the rejections, the principal will offer the same or similar wages to both agents.

A total of 288 participants took part in the experiment. The experimental sessions were conducted in the New Zealand Experimental Economics Laboratory (NZEEL) at the University of Canterbury. The participants were recruited using the online database system ORSEE (Greiner, 2004). The number of participants in a session varied from 27 to 36. All sessions were run under a single-blind social distance protocol. The treatments were implemented in an across-subject design.[18] The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

Upon entering the laboratory, participants were asked to sit in a cubicle. At the beginning of each session instructions were handed out, as well as projected onto a screen and read aloud. Any questions arising were answered in private. Each participant was randomly paired with two other participants in the laboratory to form an anonymous group of three persons, i.e. a principal, agent A and agent B, later assigned to slots. All agents then completed the quiz, which consisted of 20 questions; each question had one correct answer. The agents had 10 minutes to complete the quiz. After the quiz, all participants proceeded to the decision-making part of the experiment, which was run using the strategy method (Brandts and Charness, 2011; Selten, 1967). The participants entered their decisions and upon the completion of the experiment, were asked to complete a questionnaire for which they were paid 5 NZD. The earnings were converted

[18] One participant accidently participated two times in this experiment; so his group's observations were excluded from the data.

from francs into New Zealand Dollars at exchange rate 1 franc = 0.1 NZD. The participants played the game only once as opposed to repeatedly, in order to avoid different learning and information updating possibilities. The participants were then called one by one to receive their payment in private in the payment room in the back of the laboratory. On average, a session lasted 50 minutes including the payment and participants earned on average 12.80 NZD.[19]

[19] For reference, at the time of the experiment this was approximately 11 USD and the adult minimum wage in New Zealand was 14.25 NZD per hour.

2.4 Experimental Results

2.4.1 The Agents

The relatively high MAOs observed in the data indicate that agents are concerned with their wage relative to the principals. The strength of the principal-agent comparison may divert agents' attention away from comparisons with the paired agents or limit how large the between-agent effect could be (Knez and Camerer, 1995). Table 2.2 summarizes the averages and standard deviations of the R100 and R200 agents' MAO as well as number of observations in each treatment.

Table 2.2 The R100 and R200 Agents' MAO

Treatment	Role	Avg. min acceptable offer [francs]	Std. deviation [francs]	Number of observations
1	R100	34.258	16.452	31
	R200	65.355	34.889	31
2	R100	36.906	21.349	32
	R200	57.656	38.333	32
3	R100	41.938	21.268	32
	R200	66.438	37.431	32

I test for the differences in the agent's MAO and the paired agent's wage differentials between the treatments. Table 2.3 summarizes the results of the Mann-Whitney tests. Since the only information the agents get is the relative quiz performance and the paired agents' wage, it is interesting to test for the agents' MAO and the paired agents' wage differentials.

Table 2.3 Statistical Tests for the Agents' MAO and the Paired Agents' Wage Differentials

	The R100s' MAO and the R200s' wage differentials (Mann-Whitney p-value)	The R200s' MAO and the R100s' wage differentials (Mann-Whitney p-value)
Random assignment vs. Principal assignment	0.855 (0.393)	1.137 (0.256)
Random assignment vs. Quasi-random assignment	-0.312 (0.755)	-0.276 (0.783)
Principal assignment vs. Quasi-random assignment	-1.677 (0.094)*	-1.495 (0.135)

Hypothesis 1 states that the R100 agents' MAO and the R200 agents' wage differentials are smaller in the Principal Assignment treatment than in the Random Assignment treatment. This hypothesis is based on a conjecture that intentions and social comparisons matter and agents will state MAOs closer to the paired agents' wages when the principal does the assignment.

Result 1: The R100 agents' MAO and the R200 agents' wage differentials are not smaller in the Principal Assignment treatment than in the Random Assignment treatment.

Support for Result 1: As can be seen from Table 2.3, the Mann-Whitney test does not detect a statistically significant difference in the agents' MAO and the paired agents' wage differentials between the Random Assignment treatment and the Principal Assignment treatment (p-value=0.393). This suggests that the differentials between the

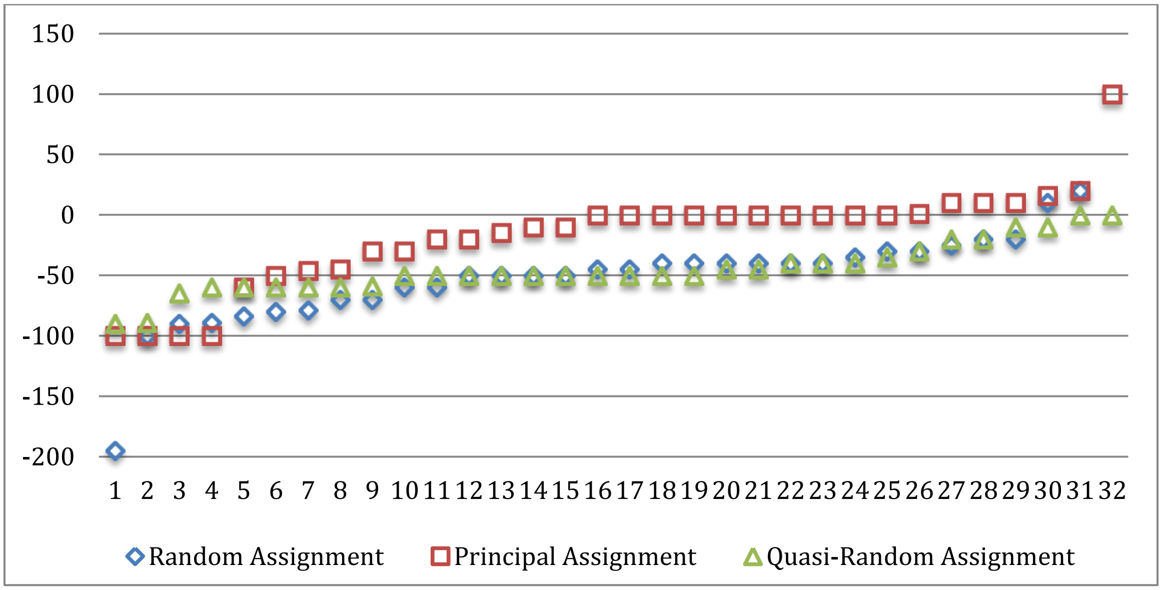
R100 agents' MAO and the R200 agent's wage are the same in the Principal Assignment treatment and in Random Assignment treatment (differentials depicted in Figure 2.2).

Hypothesis 2 states that **t**he differentials between the R100 agents' MAO and the R200 agents' wage are smaller in Quasi-Random Assignment treatment than in Random Assignment treatment.

Result 2: The differentials between the R100 agents' MAO and the R200 agents' wage are not smaller in the Quasi-Random Assignment treatment than in the Random Assignment treatment.

Support for Result 2: There is no statistically significant difference in the R100 agents' MAO and the R200 agents' wage differentials between the Quasi-Random Assignment treatment and the Random Assignment treatment (p-value=0.755). There is, however, a statistically significant difference in the R100 agents' MAO and the R200 agents' wage differentials between the Quasi-Random Assignment treatment and the Principal Assignment treatment (p-value=0.094). In the Principal Assignment treatment this differentials are smaller suggesting that agents compare their wages with the paired agents' wages.

Figure 2.2 The R100 Agents' MAO and the R200 Agents' Wage Differentials



Hypothesis 3 states that the R200 agents' MAO and the R100 agents' wage differentials are smaller in the Principal Assignment treatment than in the Random Assignment treatment. This hypothesis, again, is based on a conjecture that agents compare their wages with their paired agents' wages.

Result 3: The R200 agents' MAO and the R100 agents' wage differentials are not smaller in the Principal Assignment treatment than in the Random Assignment treatment.

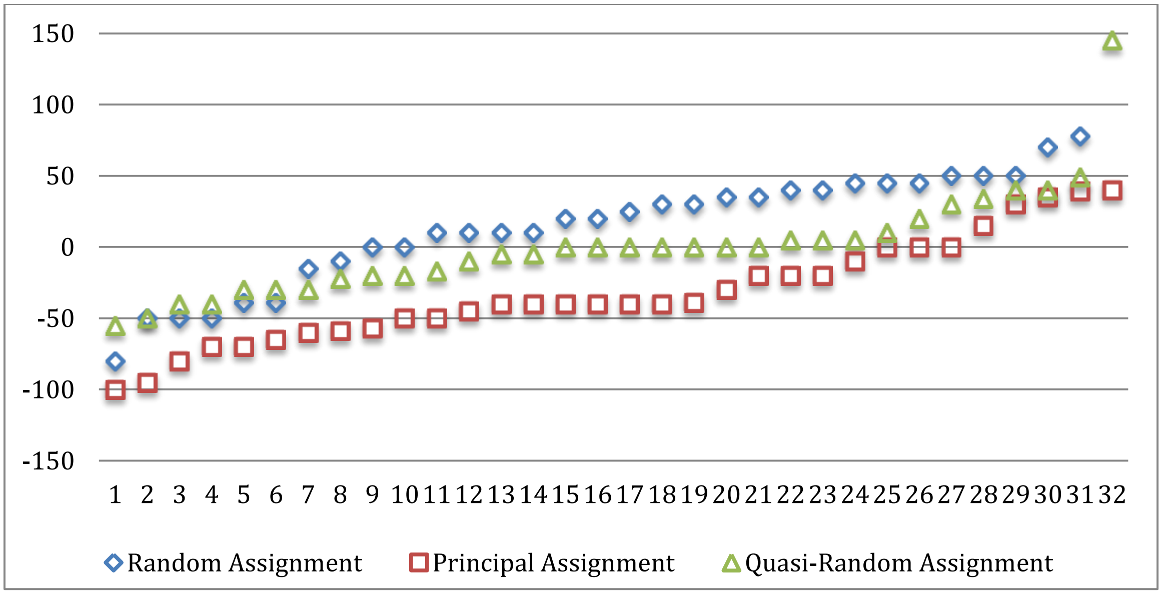
Support for Result 3: The Mann-Whitney test does not detect a statistically significant difference in the R200 agents' MAO and the R100 agents' wage differentials between the Random Assignment treatment and the Principal Assignment treatment (p-value=0.256). The R200 agents' MAO and the R100 agents' wage differentials are depicted in Figure 2.3.

Hypothesis 4 states that the differentials between R200 agents' MAO and the R100 agents' wage are smaller in the Quasi-Random Assignment treatment than in the Random Assignment treatment.

Result 4: The differentials between the R200 agents' MAO and the R100 agents' wage are not smaller in the Quasi-Random Assignment treatment than in the Random Assignment treatment.

Support for Result 4: According to the Mann-Whitney test there is no statistically significant difference in the R200 agents' MAO and the R100 agents' wage differentials between the Quasi-Random Assignment treatment and the Random Assignment treatment (p-value=0.783).

Figure 2.3 The R200 Agent's MAO and the R100 Agent's Wage Differential



For completeness, I test for the differences in the R200 and R100 agents' MAO differentials as well as the R200 agents' MAO and the R100 agents' MAO between the treatments using the Mann-Whitney test (see Table 2.4). Based on previous literature I posit that the MAOs will be higher when the principal intentionally assigns the agents to the slots as opposed to the agents being randomly assigned to the slots (Charness, 2004; Kagel, Kim, and Moser, 1996).

Table 2.4 Statistical Tests, the MAOs

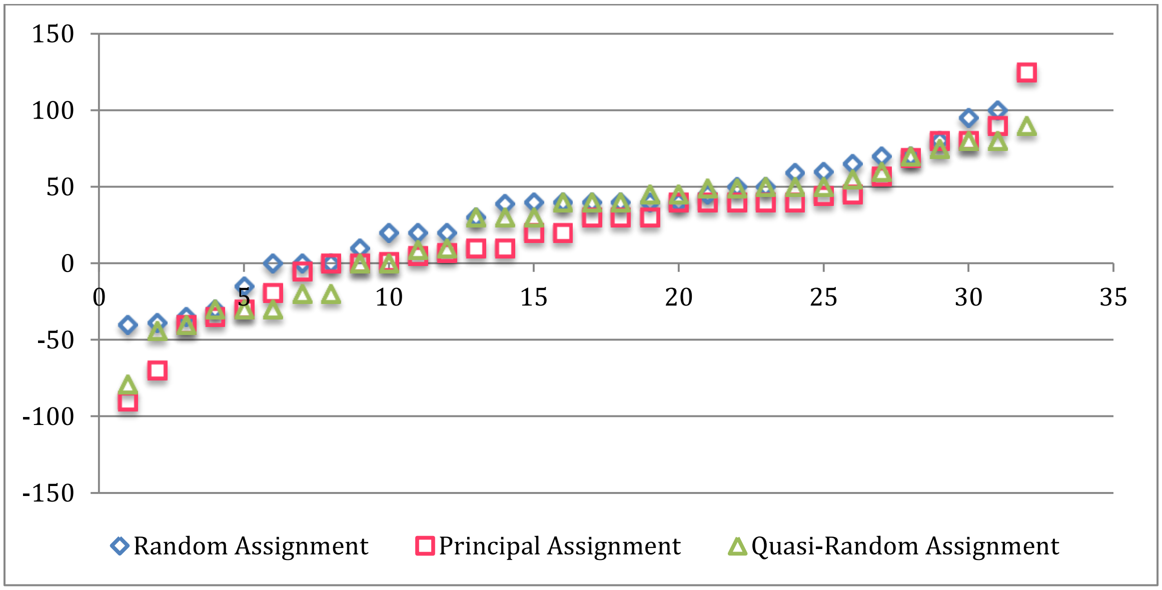
Treatment	MAO differentials (Mann-Whitney p-value)	MAOs of the R200 agents (Mann-Whitney p-value)	MAOs of the R100 agents (Mann-Whitney p-value)
Random Assignment vs. Principal Assignment	1.055 (0.291)	1.106 (0.269)	-0.552 (0.581)
Random Assignment vs. Quasi-Random Assignment	0.296 (0.767)	-0.194 (0.847)	-1.214 (0.225)
Principal Assignment vs. Quasi-Random Assignment	-0.720 (0.472)	-1.331 (0.183)	-0.521 (0.603)

The R200 agents' MAO and the R100 agents' MAO differentials are depicted in Figure 2.4. If intentions and social comparisons play a role the differentials between R200 agents' MAO and R100 agents' MAO will be higher in the Quasi-Random Assignment treatment as opposed to the Random Assignment treatment. In the Quasi-Random treatment the principals do not know the relative quiz performance and so the assignment to the slot is random from the principals' point of view and might result to more "unfair" assignments. The agents might thus react to this unfair assignment by stating higher MAOs in the Quasi-Random Assignment treatment than in the Random Assignment treatment. As can be seen from Table 2.4, the Mann-Whitney test does not detect a statistically significant difference in the R200 agents' MAO and R100 agents' MAO differentials between the Random Assignment treatment and the Principal Assignment treatment (p-value=0.291). I find no evidence of the R200 agents' MAO and R100 agents' MAO differentials to be statistically significantly different between the Quasi-Random Assignment treatment and the Random Assignment treatment (p-value=0.767). The Mann-Whitney test also does not detect a statistically significant difference in the R200 agents' MAO and R100 agents' MAO differentials between the Principal Assignment treatment and the Quasi-Random Assignment treatment (p-value=0.472).

I also tested for the differences in the R200 agents' MAO across the treatments. The Mann-Whitney test does not detect a statistically significant difference in the R200 agents' MAO between the Random Assignment treatment and the Principal Assignment treatment (p-value=0.269) or between the Random Assignment treatment and Quasi-Random Assignment treatment (p-value=0.847) or between the Principal Assignment and the Quasi-Random Assignment treatment (p-value=0.183).

The test results for the differences in the R100 agents' MAO are also reported in Table 2.4. As in the previous case, the Mann-Whitney test does not detect a statistically significant difference between the Random Assignment treatment and the Principal Assignment treatment (p-values 0.581) or the Random Assignment treatment and the Quasi-Random Assignment treatment (p-value=0.225) or the Principal Assignment treatments and the Quasi-Random treatment (p-value=0.603).

Figure 2.4 The R200 Agents' MAO And The R100 Agents' MAO Differentials



In the Random Assignment treatment, 13 out of 62 wages were rejected (21%), in the Principal Assignment treatment 13 out of 64 wages were rejected (20%) and in the Quasi-Random Assignment treatment, in total eleven out of 64 wages were rejected (17%), which can be seen in Table 2.6. The Mann-Whitney test does not detect a statistically significant difference in the rejection rate of the R200 agents (p-value= 0.716) and of the R100 agents (0.786) between the Random Assignment treatment and the Principal Assignment treatment.[20]

[20] The Mann-Whitney test shows that there is no statistically significant difference in the rejection rate of R200 agents (p-value= 0.716) and of R100 agents (0.683) between the Random Assignment treatment and the Quasi-Random Assignment treatment and also no statistically significant difference in the rejection rate of R200 agents (p-value= 1.000) and of R100 agents (0.495) between the Principal Assignment treatment and the Quasi-Random Assignment treatment.

2.4.2 The Principals

The summary of wages offered to the R200 and R100 agents is presented in Table 2.5. In the Random Assignment treatment the maximum wage offered to the R100 agent is 100 francs (there were 4 wage offers of 100 francs) and to the R200 agent 200 francs. On average, in the Random Assignment treatment the principals offered wages of 51.97 and 87.06 francs to the R100 agent and the R200 agent, respectively. In the Principal Assignment treatment the principals offered wages of 55.00 and 90.47 francs to the R100 agent and the R200 agent, respectively. In the Quasi-Random Assignment treatment the principals offered wages of 52.03 and 86.88 francs to the R100 agent and the R200 agent, respectively. Five out of 31 principals (16%) offered an equal split of the relevant pie (i.e. offering the R100 agent 50 and the R200 agent 100) to both agents in the Random Assignment treatment, 10 out of 32 principals (31%) in the Principal Assignment treatments and 11 out of 32 principals (34%) in the Quasi-Random Assignment treatment.

Table 2.5 Principals' Behaviour, Wages

	Random Assignment		**Principal Assignment**		**Quasi-Random Assignment**	
	To the R100 agent	To the R200 agent	To the R100 agent	To the R200 agent	To the R100 agent	To the R200 agent
Mean	51.97	87.06	55.00	90.47	52.03	86.88
Median	50.00	90.00	50.00	100.00	50.00	90.00
Std. deviation	22.76	30.64	15.43	18.32	14.91	18.22
Minimum	2	20	25	30	10	10
Maximum	100	200 (120*)	100	115	90	100

*The second highest wage in parentheses

The distribution of the wages and the rejection rates are shown in Table 2.6. The modal wage of the R200 agents as well as of the R100 agents is between 41-50% for all three treatments. No more than 4.8% of the wages are below 20% of the pie. This is in line with general observation made by Camerer (2003). Hence, the sub-game perfect equilibrium of very low wages is strongly rejected, consistent with previous literature.

Table 2.6 Wage Distributions and the Rejection Rates

	The Random Assignment Treatment				The Principal Assignment Treatment				The Quasi-Random Assignment Treatment			
Wage range	**Wages (percent)**		**Rejected (percent)**		**Wages (percent)**		**Rejected (percent)**		**Wages (percent)**		**Rejected (percent)**	
	R200	**R100**	**R200**	**R100**	**R200**	**R100**	**R200**	**R100**	**R200**	**R100**	**R200**	**R100**
0-10%	1 (3.2%)	1 (3.2%)	1 (100%)	1 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.1%)	1 (3.1%)	1 (100%)	0 (0%)
11-20%	1 (3.2%)	0 (0%)	1 (100%)	0 (0%)	1 (3.1%)	0 (0%)	1 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
21-30%	3 (9.7%)	5 (16.1%)	1 (33.3%)	2 (40%)	2 (6.3%)	2 (6.3%)	1 (50%)	1 (50%)	2 (6.3%)	0 (0%)	1 (50%)	0 (0%)
31-40%	8 (25.8%)	2 (6.5%)	3 (37.5%)	1 (50%)	4 (12.5%)	1 (3.1%)	3 (75%)	1 (100%)	6 (18.8%)	4 (12.5%)	3 (50%)	2 (50%)
41-50%	15 (48.4%)	16 (51.6%)	2 (13.3%)	1 (6.3%)	22 (68.8%)	19 (59.4%)	2 (9.1%)	4 (18.2%)	23 (71.9%)	19 (59.4%)	2 (8.7%)	2 (10.5%)
51-60%	2 (6.5%)	0 (0%)	0 (0%)	0 (0%)	3 (9.4%)	3 (9.4%)	0 (0%)	0 (0%)	0 (0%)	4 (12.5%)	0 (0%)	0 (0%)
61-70%	0 (0%)	3 (9.7%)	0 (0%)	0 (0%)	0 (0%)	2 (6.3%)	0 (0%)	0 (0%)	0 (0%)	1 (3.1%)	0 (0%)	0 (0%)
71-80%	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (9.4%)	0 (0%)	0 (0%)	0 (0%)	1 (3.1%)	0 (0%)	0 (0%)
81-90%	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (3.1%)	0 (0%)	0 (0%)	0 (0%)	2 (6.3%)	0 (0%)	0 (0%)
91-100%	1 (3.2%)	4 (12.9%)	0 (0%)	0 (0%)	0 (0%)	1 (3.1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

There is a clear pattern of a higher rate of rejection as the wage decreases. The R200 agents rejected eight wages in Random Assignment; seven in the Principal Assignment treatment and seven in the Quasi-Random Assignment treatment with average rejected wage of 63.13; 71.14 and 70.00 francs, respectively. R100 agents rejected five, six and four wages with average rejected wage of 29.40; 44.17 and 43.75 francs in the Random Assignment treatment, the Principal Assignment treatment and the Quasi-Random Assignment treatment, respectively. There were five observations in which the principal offered a wage of 100 francs (the entire pie) to the R100 agent (Figure 2.5) and one observation where the principal offered a wage of 200 francs to the R200 agent (Figure 2.6). Since there is a total amount of 300 francs to be divided among three people, offering 100 francs to both agents, results in an equal wage for all three participants.

Figure 2.5 Wage As A Percentage Of The Pie Offered To The R200 Agents

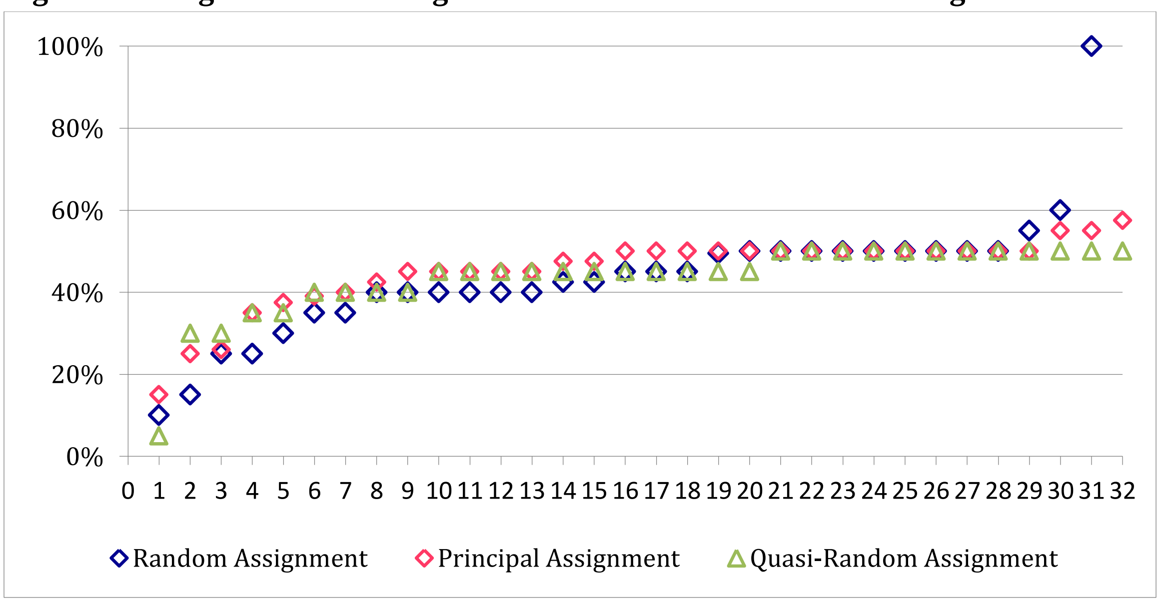
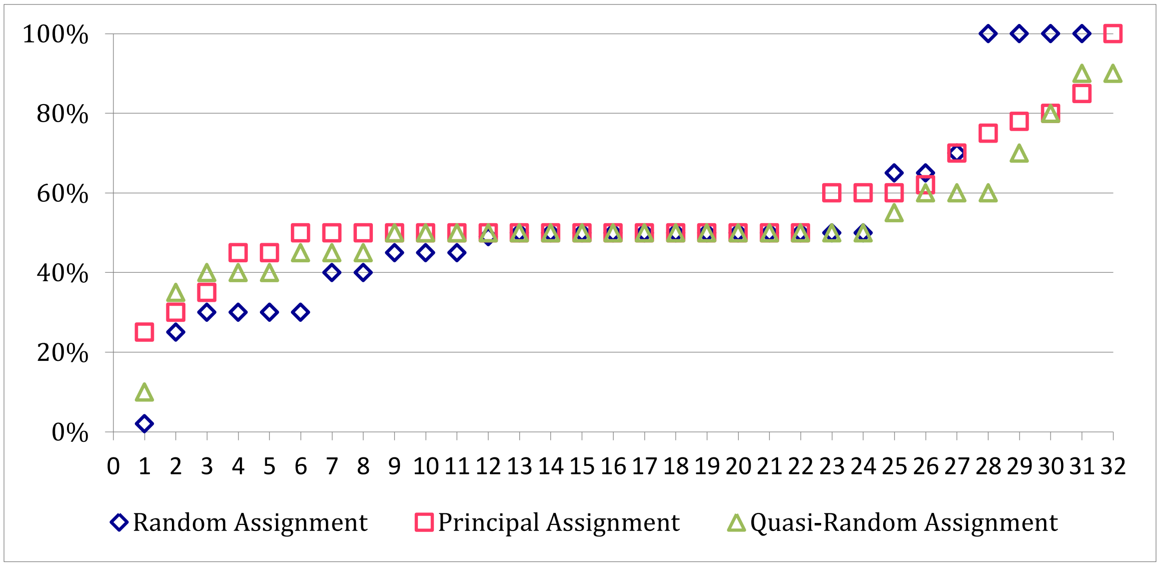


Figure 2.6 Wage As A Percentage Of The Pie Offered To The R100 Agents



Hypothesis 3 states that the R100 agents' wage and the R200 agents' wage differentials will be higher in the Random Assignment treatment than in the Principal Assignment treatment. If the agents are sensitive to the assignment made by the principal, they will be more sensitive to wages (they will report higher MAOs) in the treatment where agents are assigned to slots by the principal. On one hand, the principals, who anticipate this and want to maximize their payoff, might want to offer similar wages to the agents to avoid their rejections when the agents compare their own wage to that of the paired agents. As a result the principal will offer similar wages to the R100 and R200 agents. On the other hand, I expect the principal to anticipate that the agent is less sensitive to the paired agent's wage when the assignment to the slots is random as opposed to the principal making the assignment. The principal will thus offer higher wages to the R200 agents than to the R100 agents.

Result 5: I find no evidence of wage compression in the Principal Assignment treatment as opposed to Random Assignment.

Support for Result 5: I find no statistically significant difference in the R200 and R100 agents' wage differentials between Random Assignment and the Principal Assignment treatment (Mann-Whitney p-value=0.494). I also do not find a statistically significant result when it comes to wage differentials between the Random Assignment Treatment and the Quasi-Random Assignment Treatment (p-value=0.731) and the Principal Assignment Treatment and Quasi-Random Assignment Treatment (p-value=0.694).[21]

Figure 2.7 shows the wages as a percentage of the pie from the principal's point of view. If the principal offered 50% of the pie to the R100 agents and also 50% of the pie to the R200 agents, the wages would be lying on a 45-degree line (projected as red line in the picture). It can be seen from the Figure 2.7 that there was only one wage to the R200 agent that exceeded 60% of the pie. However, there are a few wages ranging from 60-100% among the wages to the R100 agents. It appears that the principals are willing to offer bigger share of the pie to the R100 agents than to the R200 agents.

[21] Since the wage of 200 francs to the R200 agent in the Random Assignment treatment can be considered an outlier, I ran the tests excluding this observation as well. There is still no statistical difference in the wages offered to the R100 and the R200 agents between the Random Assignment treatment and Principal Assignment treatment (Mann-Whitney p-value=0.276); the Random Assignment treatment and the Quasi-Random Assignment treatment (p-value=0.560); and the Principal Assignment treatment and the Quasi-Random Assignment treatment (p-value=0.615).

Figure 2.7 Wages offered to both agents [%]

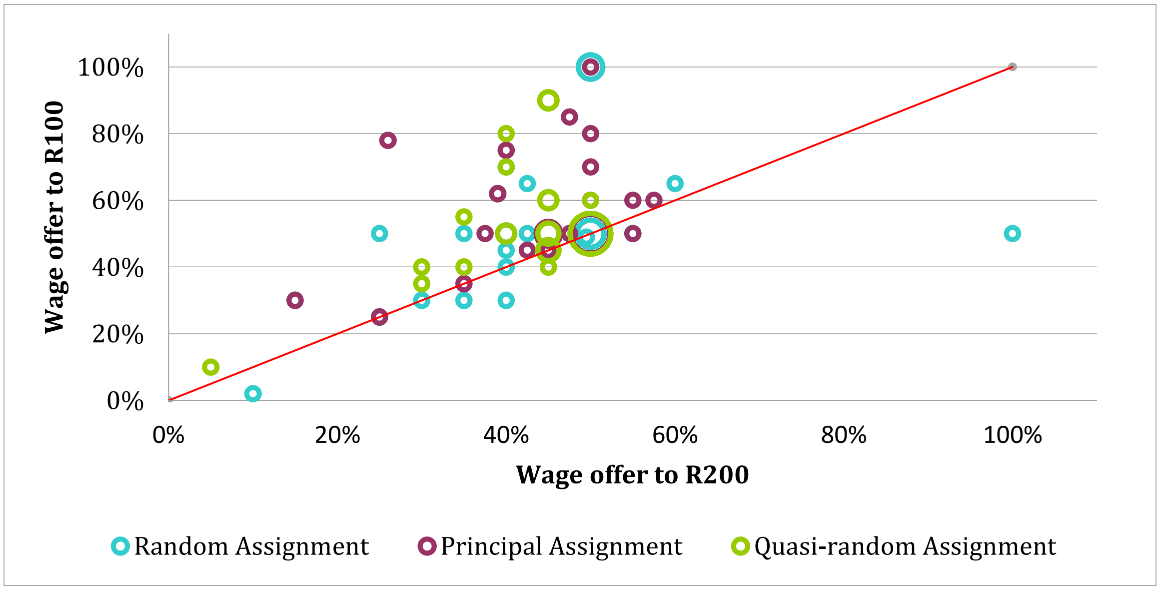


Table 2.7 The Mann-Whitney Test Result after Discarding Unfair Assignments from the Principal Assignment Treatment

Treatment	Wage differentials (Mann-Whitney p-value)
Random Assignment vs. Principal Assignment	-0.863 (0.388)
Random Assignment vs. Quasi-Random Assignment	-
Principal Assignment vs. Quasi-Random Assignment	0.586 (0.558)

After discarding the wages in which the principal assigned the agent with higher relative quiz performance to the low slots, I tested if there is a statistically significant difference in the R100 agents' and R200 agents' wage differentials between treatments Random Assignment and the Principal Assignment treatment; and the Principal Assignment treatment and the Quasi-Random Assignment treatment. Notice that only in the Principal Assignment treatment the principal knew the relative quiz performance and decided on the assignment of the agents to the slots. The Mann-Whitney test does not detect a statistically significant difference in the R100 and R200 agents' wage differentials (see Table 2.7).

2.5 Demographics

2.5.1 The Principals

As demographic data were elicited in the post-experiment questionnaire I can test if any of these are influencing the principal's decision on wage offered to the R200 agent and the R100 agent. All variables that were potentially relevant were regressed and the results from an OLS regression are reported in Table 2.8 below.

THE RANDOM ASSIGNMENT TREATMENT is omitted from the regression to use it as the comparison basis for treatment effects, and to avoid perfect correlation. THE PAIRED AGENT'S WAGE is the wage offered to the R100 agent if the dependent variable is the R200 agent's wage and to the R200 agent if the dependent variable is the R100 agent's wage. This variable is statistically significant in both regressions. MALE is a dummy variable that is 1 when a subject reported their gender as male and 0 if female. NON NZ is a dummy variable that is 1 when subjects nominated a nationality of a country that was not New Zealand, out of this group approximately 55.6% of the principals and 61% of the agents nominated an Asian nationality, and approximately 16.7% of the principals and 15.3% of the agents nominated a European nationality. ECON is a dummy variable that is 1 when subjects nominated an area of study that was 'Economics', or 'Business Economics'.

Table 2.8 Demographics Analysis, The Principals

	OLS Coefficient (p-value)	
	Dependent variable the R200 agent's wage	**Dependent variable the R100 agent's wage**
THE PRINCIPAL ASSIGNMENT TREATMENT	3.229 (0.554)	2.562 (0.554)
THE QUASI-RANDOM ASSIGNMENT TREATMENT	2.234 (0.673)	0.046 (0.991)
THE PAIRED AGENT'S WAGE	0.502*** (0.001)	0.316*** (0.001)
MALE	-8.969** (0.043)	1.440 (0.686)
AGE	1.281** (0.013)	-0.157 (0.707)
SIBLINGS	1.507 (0.315)	-0.478 (0.689)
ECON	-4.439 (0.313)	-0.805 (0.818)
NON NZ	0.000 (1.000)	-2.057 (0.585)
RELATIVE INCOME	1.404 (0.548)	1.495 (0.419)
CITY SIZE	0.147 (0.952)	-2.210 (0.249)
LIVE WITH OTHERS	0.050 (0.831)	0.429** (0.018)
MONEY	0.002 (0.682)	0.001 (0.865)
FINANCE STUDY	-0.045 (0.488)	-0.083 (0.101)
NO. OF PEOPLE KNOWN	3.004 (0.301)	-2.802 (0.223)
CONSTANT	30.839 (0.056)	35.109 (0.005)

Run on StataSE 12.0. *, **, *** refer to statistical significance at the 10%, 5% and 1% levels, respectively.

LIVE WITH OTHERS is the number of people that currently live in the subject's household, MONEY is the amount of dollars that subjects nominated as their monthly non-accommodation budget, FINANCE STUDY is the proportion that subjects nominated as the fraction of their monthly budget that they fund themselves. NO. OF PEOPLE KNOWN is the number subject knows in the session. RELATIVE INCOME is the income of subjects' parents in comparison to other families in New Zealand when subjects were 16 years of age. The larger the income of the family they state the larger the variable RELATIVE INCOME. Subjects had five possibilities to choose from, i.e. far below average, below average, average, above average, far above average. CITY SIZE is the size of the community where the subject has lived the most time of their life. Subjects could choose from the following four options: up to 2 000 inhabitants, 2 000 to 10 000 inhabitants, 10 000 to 100 000 inhabitants, more than 100 000 inhabitants. The more inhabitants they state the higher is the variable CITY SIZE. MONEY and FINANCE STUDY were included as measures to control for income.

The coefficients for THE PRINCIPAL ASSIGNMENT TREATMENT and THE QUASI-RANDOM ASSIGNMENT TREATMENT confirm that there is no statistically significant difference in the R200 and R100 agents' wage differentials compared to THE RANDOM ASSIGNMENT TREATMENT. Interestingly, the variable MALE and AGE are statistically significant at 5% level in the R200 agent's wage but not in the R100 agent's wage. The variable ECON was used to control for the behaviour of economics students. In this game it would be predicted that economics students will offer a wage of 0. Even though the coefficient is not statistically significant, the sign is in the correct direction.

An interesting result of this analysis of demographics is the statistically significant coefficient (the R100 agent's wage) related to the variable LIVE WITH OTHERS. This suggests that the more people that live in the same household the higher wage the subjects tend to offer to the R100 agents.

2.5.2 The Agents

Demographic data were also elicited from the agents. Therefore, I can test if any of the demographic data affected the agents' decision on their minimum acceptable offer. The OLS regression (presented in Table 2.9) was run on any variables that could have an impact on the agents' behaviour.

As previously, with the demographics that were elicited for the principals, the treatments do not have an effect on the minimum acceptable offers. The variable THE PAIRED AGENT'S WAGE is not statistically significant, which confirms the result that the paired agent's wage does not influence the agent's minimum acceptable offer. The only statistically significant variable in the regression is R200 AGENT, meaning that the agents assigned to the high slot stated higher minimum acceptable offers than the agents assigned to the low slots. It appears from the results that demographics are not significantly biasing results in this experiment.

Table 2.9 Demographics Analysis, The Agents

	OLS Coefficient (p-value) Minimum Acceptable Offer
THE PRINCIPAL ASSIGNMENT TREATMENT	-4.092 (0.492)
THE QUASI-RANDOM ASSIGNMENT TREATMENT	2.855 (0.633)
R200 AGENT	14.355** (0.033)
THE PAIRED AGENT'S WAGE	0.117 (0.356)
MALE	2.391 (0.632)
AGE	0.316 (0.510)
SIBLINGS	-0.760 (0.645)
ECON	-1.178 (0.815)
NON NZ	-2.074 (0.710)
RELATIVE INCOME	-2.688 (0.301)
CITY SIZE	-0.806 (0.748)
LIVE WITH OTHERS	0.147 (0.669)
MONEY	0.006 (0.273)
FINANCE STUDY	0.051 (0.447)
NO. OF PEOPLE KNOWN	3.421 (0.213)
CONSTANT	32.057 (0.149)

Run on StataSE 12.0. *, **, *** refer to statistical significance at the 10%, 5% and 1% levels, respectively.

2.6 Conclusion

Previous literature provides evidence that agents care about their own wages, the wages of other workers, and about the intentions of the principal. In the current experiment I test whether agents are sensitive to the paired agents' wage, to the process of the assignment to slots, and whether the principal compresses wages in treatments where the assignment might seem less fair. The results from the current experiment cast doubt on the notion that agents' concerns with the paired agents' wages might explain pay policies such as wage compression or wage secrecy. The managers of the firms decide to compress wages (Akerlof and Yellen, 1990) when there is evidence of wage comparisons among the agents, which can jeopardize firms' monetary performance. The current experiment does not provide evidence on principals compressing wages in a treatment with unfair assignment to slots.

Fairness norms are highly relevant for work-related decisions in organizations and on the labour market (e.g. Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999; Gächter and Fehr, 2002; Kahneman, Knetsch, and Thaler, 1986; Konow, 1996; Rabin, 1993; Scott, 2003). Some theories of fairness, in particular inequity aversion by Fehr and Schmidt (1999) and Bolton and Ockenfels (2000), can be used to help explain subject behaviour. According to inequity aversion people resist inequitable outcomes; i.e., they are willing to give up some material payoff to move in the direction of more equitable outcomes. In a game with more than two players Fehr and Schmidt (1999) assume that a player compares himself to each of the other players separately. This implies that his behaviour towards another player depends on the income difference towards this player. Behaviour of some subjects can be explained by the Fehr and Schmidt theory of inequity as 25.3% (24 out of 95 principals) behaved consistently with this theory (offering exactly 50 to the R100 agents and exactly 100 to the R200 agents). Bolton and Ockenfels (2000), on the other hand, assume that the decision maker is not concerned about each individual opponent but only about the average income of all players.[22] Based on this theory it is possible that the principals were trying to be fair by offering a "fair share", i.e. exactly 100 francs for each of the agents and 100 francs to keep for themselves. As can be seen from the data 5% (5 out of 95 principals) behaved consistently with this theory.

The complex settings of firms, such as multiple hierarchies and relationships, make it hard to understand aspects of fairness behaviour as multiple reference points might be behind the decisions of people (Alewell and Nicklisch, 2006). Firms are usually structured into a hierarchy where many agents belong to the same hierarchical layer and are concerned about horizontal fairness (Güth, Königstein, Kovács, and Zala-Mezõ, 2001). Several recent papers provide laboratory evidence on the role of horizontal social comparisons for agents' effort choices in a gift exchange environment (e.g. Clark, Masclet, and Villeval, 2006; Gächter and Thöni, 2010). However, between-agent comparisons are not the only comparisons being made in a workplace environment. Another important fairness aspect to have in mind is vertical fairness, i.e. the relationship between the outcome distributions among principals and agents, and agents' acceptance thresholds (Güth, Königstein, Kovács, and Zala-Mezõ, 2001). Agents might consider their principals' earnings as a reference point, based on which they judge if they have been treated fairly (Adams, 1965; Deutsch, 1975; Leventhal, 1976b). Vertical fairness concerns may lead firms to pay more generous wages, leading to particular wage

[22] Their model assumes that subjects like the average earnings to be as close as possible to their own earnings (Engelmann and Strobel, 2000).

dynamics and possibly explaining inter-industry wage differentials (Krueger and Summers, 1988). Economic experiments have also provided evidence of the importance of vertical fairness. Croson (1996) and Straub and Murnighan (1995) found that agents accepted lower wages when they did not know how much the principal earned. Cabrales and Charness (2000) find that the principal's payoff was important for agent behaviour. Charness and Kuhn (2007) did not find that agents' effort is affected by other agent's wage, questioning why firms adopt policies of wage secrecy and wage compression.

The introduction of multiple possible reference points for fairness judgements results in a self-serving bias (Babcock, Wang, and Loewenstein, 1996) or egocentric selections between these different points of fairness judgements for each player. In the ultimatum game with more than two players, for example, principals select a fairness standard which leads to relatively low wages, agents focus on reference points which lead to high wages, and therefore reject lower wages more frequently than in the simple ultimatum game. Moreover, even for those agents who apply social comparisons between agents, the strength of this motive seems to somewhat depend on the size of the wage received by the principal and the other agent (Alewell and Nicklisch, 2006). This could limit how large the between-agent effect could be (Knez and Camerer, 1995). In the current experiment there is no statistically significant difference in the agents' MAO and the paired agents' wage differentials across treatments (except for the R100 agents' MAO and the R200 agents' wage differentials being statistically significantly smaller in the Principal Assignment treatment than in the Quasi-Random Assignment treatment). The original conjecture that these differentials will be smaller in the Principal Assignment treatment and the Quasi-Random Assignment treatment than in the Random Assignment treatment is not supported. There are relatively high MAOs observed in the current experiment, which indicates that agents might be concerned with their wages relative to the principal's wage. Agents' comparisons of their earnings with the principals' earnings thus cause rejection of low wages (by agents stating higher MAOs). Instead of horizontal fairness, vertical fairness could be the main consideration. This suggests that horizontal fairness may only matter when fairness on the vertical level is satisfied. Further research on separating principal-agent comparisons from agent-agent comparisons is therefore needed to address social comparisons.

Another reason why there was no evidence of social comparisons in the current experiment might be that even though the agents were comparing their wages with the paired agents' wages, they were reluctant to take an action in response because of loss aversion (Kahneman and Tversky, 1984). Even though they were comparing themselves with the paired agents and it made them feel unfair, it was not unfair enough for them to risk losing their wage. In other words, even if an agent felt he has been treated unfairly, he did not necessarily reject the wage by stating a higher minimum acceptable offer. This is despite the fact that there is a large body of experimental evidence on importance of the social comparisons (e.g. Clark and Senik, 2010; Falk and Knell, 2004; Gächter and Thöni, 2010). The data from the current experiment suggests that social comparisons might be of secondary importance in labour relationships or cannot be generalised, i.e. it has different implications in different industries. Social comparisons might be more easily detectable in bargaining, where they are often used as a bargaining strategy to justify one's entitlement, which might be another fruitful avenue for further research.

Previous research on social comparisons has implications for the world outside of the laboratory. It suggests that there is a reason why wages are a secret in most companies (Danziger and Katz, 1997; Lawler, 1990) and most of the time the principal's earnings (or in real world terms, the firm's earnings) are not publicly available.[23] From the current experiment it appears that horizontal fairness doesn't matter when vertical fairness isn't satisfied. Vertical fairness can be satisfied by wage (profit) secrecy, which then can lead to the importance of horizontal fairness as observed in the field. If vertical fairness is not satisfied, principals do not need to worry about social comparisons between agents and can offer different wages to different agents. It also appears that the fairness of job assignments does not matter. This would mean that agents with marginally different skills do not consider the intentions of their principal when choosing whether to provide effort, shirk or quit the job when treated unfairly. Future research is needed to determine the relationship of vertical and horizontal fairness and its potential link to social comparisons if one of these is satisfied and the other is not.

[23] If the firm is a publicly traded company their financial information is made available, however, it is done in a complicated manner and they average person doesn't really understand these.

Appendix A. Subject Instructions

The Random Assignment Treatment

INSTRUCTIONS

No talking allowed

Thank you for coming. The purpose of this session is to study how people make decisions in a particular situation. From now until the end of the session, unauthorized communication of any nature with other participants is prohibited. If you violate this rule we will have to exclude you from the experiment and from all payments. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Earnings

Every participant will have an opportunity to earn money in the experiment. Your final experimental earnings will depend on your decisions and on the decisions of others. It is therefore very important that you read these instructions carefully. The payoffs will be denoted in experimental currency referred to as francs. Upon completion of the experiment, all francs will be exchanged into dollars using the following exchange rate:

1 franc = $0.1

Notice that the more francs you earn, the more dollars you will receive. All the money will be paid to you privately in cash at the end of the experiment.

Group Assignments

You will be randomly paired with two other participants to form a group of three persons. No one will learn the identity of the persons (s)he is paired with. Each person in the group will be assigned to serve as either “a Proposer” or “a Recipient”. Each group consists of one Proposer and two Recipients: Recipient A and Recipient B. The computer randomly determines whether you will be a Proposer, Recipient A or Recipient B and will inform you about your assignment at the beginning of the experiment. You have a 1/3 chance of becoming a Proposer, a 1/3 chance of becoming Recipient A, and a 1/3 chance of becoming Recipient B.

General Knowledge Quiz

In the first part of the experiment the Recipients will be asked to complete a general knowledge quiz. Each Recipient will be asked to answer the same set of 20 questions in the same order. Each question has one correct answer. The Recipients will have 10 minutes to answer all 20 questions. Remaining unanswered questions count as incorrect answers.

Within each group, the two Recipients will be ranked based on their quiz scores. If both Recipients have the same score in the quiz, the Recipient who completed the quiz more quickly will be ranked higher. The Proposer and both Recipients will be informed about which of the Recipients ranked higher.

While Recipients complete the quiz, we ask all Proposers to wait patiently and quietly. Please do not use the computer in front of you as it is set up for the experiment.

Decision-Making Part

Within each group, one randomly selected Recipient will be assigned the R-200 role, and the other Recipient will be assigned the R-100 role. Each Recipient has a 1in 2 (i.e. 50%) chance of being assigned the R-200 role, and also 1 in 2 (i.e. 50%) chance of being assigned the R-100 role.

The Proposer and the R-200 Recipient will receive a sum of 200 francs to be divided between themselves. Separately, the Proposer and the R-100 Recipient will receive a sum of 100 francs to be divided between themselves.

The procedure for dividing each sum of money between the Proposer and each Recipient is as follows. The Proposer will choose how many francs out of 200 to offer to the R-200 Recipient and how many francs out of 100 to offer to the R-100 Recipient.

If the R-200 Recipient accepts the offer made to him/her, (s)he will receive the number of francs stated in the offer, whereas the Proposer will keep the remainder, (200 – the offer). If the R-200 Recipient rejects the offer, the 200 francs disappears and both the Proposer and the R-200 Recipient will receive nothing.

Similarly, if the R-100 Recipient accepts the offer made to him/her, (s)he will receive the number of francs stated in the offer, whereas the Proposer will keep the remainder, (100 – the offer). If the R-100 Recipient rejects the offer, the 100 francs disappears and both the Proposer and the R-100 Recipient will receive nothing.

Each Recipient will not observe the offer that the Proposer made to him/her; however, (s)he observes the offer that the Proposer made to the *other* Recipient. After observing the offer that the Proposer made to the other Recipient, each Recipient chooses a number (an integer between zero and the total sum, which is 200 francs for the R-200 Recipient and 100 francs for the R-100 Recipient). This number represents the minimum offer that (s)he is willing to accept from the Proposer, so we call this number the *Minimum Acceptable Offer*. That is, if the offer made by the Proposer turns out to be greater or equal to this number, the offer is accepted. However, if the offer is less than this number then the offer is rejected. It is important to understand that each Recipient chooses the minimum acceptable offer before (s)he comes to know his/her actual offer. The decision procedure described above will be conducted only once.

Calculation of Experimental Payoffs

If the Proposer's offer to a Recipient turns out to be greater than or equal to that Recipient's Minimum Acceptable Offer, then the offer is accepted. This means the Recipient receives the amount of the offer and the Proposer receives the remainder (i.e. the total sum minus the offer made to the Recipient).

If the Proposer's offer to a Recipient turns out to be less than that Recipient's Minimum Acceptable Offer, then the offer is rejected and the Proposer and the Recipient both receive zero francs.

Notice that each Recipient's payoff is not affected by the Proposer's offer to the other Recipient, or by whether that offer (to the other Recipient) is accepted.

A hypothetical example for demonstration purposes

Suppose that:

- *Recipient A is randomly assigned the R-200 role. Recipient B is randomly assigned the R-100 role.*
- *The Proposer offers Recipient A 80 francs (out of 200)*
- *The Proposer offers Recipient B 40 francs (out of 100)*
- *Recipient A chooses a Minimum Acceptable Offer of 60 francs*
- *Recipient B chooses a Minimum Acceptable Offer of 50 francs*

This example results in the following payoffs:

- *Recipient A:*

In this case, the Proposer offered 80 francs, which is more than 60 francs, the minimum amount Recipient A would accept.

Payoffs: The Proposer receives 200-80= 120 francs and Recipient A receives 80 francs.

- *Recipient B*:

In this case, the Proposer offered 40 francs, which is less than 50 francs, the minimum amount Recipient B would accept.

Payoffs: The Proposer receives 0 francs and Recipient B receives 0 francs.

- *Proposer:*

From above, the Proposer receives 120 francs from his/her interaction with Recipient A (the remainder of the 200 francs), and receives 0 francs from his/her interaction with Recipient B (as the Proposer's offer was rejected). Thus in total the Proposer receives 0+120=120 francs.

Summary

If you are randomly selected to be the Proposer, you will have to choose an offer for each of the two Recipients. If you are randomly selected to be a Recipient, you will learn about the other Recipient's offer and will then have to state the minimum offer you are willing to accept.

Are there any questions?

The Principal Assignment Treatment

INSTRUCTIONS

No talking allowed

Thank you for coming. The purpose of this session is to study how people make decisions in a particular situation. From now until the end of the session, unauthorized communication of any nature with other participants is prohibited. If you violate this rule we will have to exclude you from the experiment and from all payments. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Earnings

Every participant will have an opportunity to earn money in the experiment. Your final experimental earnings will depend on your decisions and on the decisions of others. It is therefore very important that you read these instructions carefully. The payoffs will be denoted in experimental currency referred to as francs. Upon completion of the experiment, all francs will be exchanged into dollars using the following exchange rate:

1 franc = $0.1

Notice that the more francs you earn, the more dollars you will receive. All the money will be paid to you privately in cash at the end of the experiment.

Group Assignments

You will be randomly paired with two other participants to form a group of three persons. No one will learn the identity of the persons (s)he is paired with. Each person in the group will be assigned to serve as either "a Proposer" or "a Recipient". Each group consists of one Proposer and two Recipients: Recipient A and Recipient B. The computer randomly determines whether you will be a Proposer, Recipient A or Recipient B and will inform you about your assignment at the beginning of the experiment. You have a 1/3 chance of becoming a Proposer, a 1/3 chance of becoming Recipient A, and a 1/3 chance of becoming Recipient B.

General Knowledge Quiz

In the first part of the experiment the Recipients will be asked to complete a general knowledge quiz. Each Recipient will be asked to answer the same set of 20 questions in the same order. Each question has one correct answer. The Recipients will have 10 minutes to answer all 20 questions. Remaining unanswered questions count as incorrect answers.

Within each group, the two <u>Recipients</u> will be ranked based on their quiz scores. If both Recipients have the same score in the quiz, the Recipient who completed the quiz more quickly will be ranked higher. The Proposer and both Recipients will be informed about which of the Recipients ranked higher.

While Recipients complete the quiz, we ask all Proposers to wait patiently and quietly. Please do not use the computer in front of you as it is set up for the experiment.

Decision-Making Part

Within each group, one Recipient will be assigned the R-200 role and the other Recipient will be assigned the R-100 role. The Proposer decides which Recipient is assigned the R-200 role and which the R-100 role.

The Proposer and the R-200 Recipient will receive a sum of 200 francs to be divided between themselves. Separately, the Proposer and the R-100 Recipient will receive a sum of 100 francs to be divided between themselves.

The procedure for dividing each sum of money between the Proposer and each Recipient is as follows. The Proposer will choose how many francs out of 200 to offer to the R-200 Recipient and how many francs out of 100 to offer to the R-100 Recipient.

If the R-200 Recipient accepts the offer made to him/her, (s)he will receive the number of francs stated in the offer, whereas the Proposer will keep the reminder, (200 – the offer). If the R-200 Recipient rejects the offer, the 200 francs disappears and both the Proposer and the R-200 Recipient will receive nothing.

Similarly, if the R-100 Recipient accepts the offer made to him/her, (s)he will receive the number of francs stated in the offer, whereas the Proposer will keep the remainder, (100 – the offer). If the R-100 Recipient rejects the offer, the 100 francs disappears and both the Proposer and the R-100 Recipient receive nothing.

Each Recipient will not observe the offer that the Proposer made to him/her; however, (s)he observes the offer that the Proposer made to the *other* Recipient. After observing the offer that the Proposer made to the other Recipient, each Recipient chooses a number (an integer between zero and the total sum, which is 200 francs for the R-200 Recipient and 100 francs for the R-100 Recipient). This number represents the minimum offer that (s)he is willing to accept from the Proposer, so we call this number the *Minimum Acceptable Offer*. That is, if the offer made by the Proposer turns out to be greater or equal to this number, the offer is accepted. However, if the offer is less than this number then the offer is rejected. It is important to understand that each Recipient chooses the minimum acceptable offer before (s)he comes to know his/her actual offer. The decision procedure described above will be conducted only once.

Calculation of Experimental Payoffs

If the Proposer's offer to a Recipient turns out to be greater than or equal to that Recipient's Minimum Acceptable Offer, then the offer is accepted. This means the Recipient receives the amount of the offer and the Proposer receives the remainder (i.e. the total sum minus the offer made to the Recipient).

If the Proposer's offer to a Recipient turns out to be less than that Recipient's Minimum Acceptable Offer, then the offer is rejected and the Proposer and the Recipient both receive zero francs.

Notice that each Recipient's payoff is not affected by the Proposer's offer to the other Recipient, or by whether that offer (to the other Recipient) is accepted.

A hypothetical example for demonstration purposes

Suppose that:

- Recipient A is assigned by the Proposer the R-200 role. Recipient B is assigned by the Proposer the R-100 role.
- The Proposer offers Recipient A 80 francs (out of 200)
- The Proposer offers Recipient B 40 francs (out of 100)
- Recipient A *chooses a Minimum Acceptable Offer of 60 francs*
- Recipient B *chooses a Minimum Acceptable Offer of 50 francs*

This example results in the following payoffs:

- *Recipient A:*

In this case, the Proposer offered 80 francs, which is more than 60 francs, the minimum amount Recipient A would accept.

Payoffs: The Proposer receives 200-80=120 francs and Recipient A receives 80 francs.

- *Recipient B:*

In this case, the Proposer offered 40 francs, which is less than 50 francs, the minimum amount Recipient B would accept.

Payoffs: The Proposer receives 0 francs and Recipient B receives 0 francs.

- *Proposer:*

From above, the Proposer receives 120 francs from his/her interaction with Recipient A(the remainder of 200 francs), and receives 0 francs from his/her interaction with Recipient B(as the Proposer's offer was rejected). Thus in total the Proposer receives 0+120=120 francs.

Summary

If you are randomly selected to be the Proposer you will first have to assign the R-200 role and the R-100 role to the Recipients. Then you will have to choose an offer for each of the two Recipients. If you are randomly selected to be a Recipient, you will learn about the other Recipient's offer and will then have to state the minimum offer you are willing to accept.

Are there any questions?

The Quasi-Random Assignment Treatment

INSTRUCTIONS

No talking allowed

Thank you for coming. The purpose of this session is to study how people make decisions in a particular situation. From now until the end of the session, unauthorized communication of any nature with other participants is prohibited. If you violate this rule we will have to exclude you from the experiment and from all payments. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Earnings

Every participant will have an opportunity to earn money in the experiment. Your final experimental earnings will depend on your decisions and on the decisions of others. It is therefore very important that you read these instructions carefully. The payoffs will be denoted in experimental currency referred to as francs. Upon completion of the experiment, all francs will be exchanged into dollars using the following exchange rate:

1 franc = $0.1

Notice that the more francs you earn, the more dollars you will receive. All the money will be paid to you privately in cash at the end of the experiment.

Group Assignments

You will be randomly paired with two other participants to form a group of three persons. No one will learn the identity of the persons (s)he is paired with. Each person in the group will be assigned to serve as either "a Proposer" or "a Recipient". Each group consists of one Proposer and two Recipients: Recipient A and Recipient B. The computer randomly determines whether you will be a Proposer, Recipient A or Recipient B and will inform you about your assignment at the beginning of the experiment. You have a 1/3 chance of becoming a Proposer, a 1/3 chance of becoming Recipient A, and a 1/3 chance of becoming Recipient B.

General Knowledge Quiz

In the first part of the experiment the Recipients will be asked to complete a general knowledge quiz. Each Recipient will be asked to answer the same set of 20 questions in the same order. Each question has one correct answer. The Recipients will have 10 minutes to answer all 20 questions. Remaining unanswered questions count as incorrect answers.

Within each group, the two Recipients will be ranked based on their quiz scores. If both Recipients have the same score in the quiz, the Recipient who completed the quiz more quickly will be ranked higher.

While Recipients complete the quiz, we ask all Proposers to wait patiently and quietly. Please do not use the computer in front of you as it is set up for the experiment.

Decision-Making Part

Within each group, one Recipient will be assigned the R-200 role and the other Recipient will be assigned the R-100 role. The Proposer decides which Recipient is assigned the R-200 role and which the R-100 role.

The Proposer and the R-200 Recipient will receive a sum of 200 francs to be divided between themselves. Separately, the Proposer and the R-100 Recipient will receive a sum of 100 francs to be divided between themselves.

The procedure for dividing each sum of money between the Proposer and each Recipient is as follows. The Proposer will choose how many francs out of 200 to offer to the R-200 Recipient and how many francs out of 100 to offer to the R-100 Recipient.

If the R-200 Recipient accepts the offer made to him/her, (s)he will receive the number of francs stated in the offer, whereas the Proposer will keep the reminder, (200 – the offer). If the R-200 Recipient rejects the offer, the 200 francs disappears and both the Proposer and the R-200 Recipient will receive nothing.

Similarly, if the R-100 Recipient accepts the offer made to him/her, (s)he will receive the number of francs stated in the offer, whereas the Proposer will keep the remainder, (100 – the offer). If the R-100 Recipient rejects the offer, the 100 francs disappears and both the Proposer and the R-100 Recipient receive nothing.

Each Recipient will not observe the offer that the Proposer made to him/her; however, (s)he observes the offer that the Proposer made to the *other* Recipient. After observing the offer that the Proposer made to the other Recipient, each Recipient chooses a number (an integer between zero and the total sum, which is 200 francs for the R-200 Recipient and 100 francs for the R-100 Recipient). This number represents the minimum offer that (s)he is willing to accept from the Proposer, so we call this number the *Minimum Acceptable Offer*. That is, if the offer made by the Proposer turns out to be greater or equal to this number, the offer is accepted. However, if the offer is less than this number then the offer is rejected. It is important to understand that each Recipient chooses the minimum acceptable offer before (s)he comes to know his/her actual offer. The decision procedure described above will be conducted only once.

Calculation of Experimental Payoffs

If the Proposer's offer to a Recipient turns out to be greater than or equal to that Recipient's Minimum Acceptable Offer, then the offer is accepted. This means the Recipient receives the amount of the offer and the Proposer receives the remainder (i.e. the total sum minus the offer made to the Recipient).

If the Proposer's offer to a Recipient turns out to be less than that Recipient's Minimum Acceptable Offer, then the offer is rejected and the Proposer and the Recipient both receive zero francs.

Notice that each Recipient's payoff is not affected by the Proposer's offer to the other Recipient, or by whether that offer (to the other Recipient) is accepted.

A hypothetical example for demonstration purposes

Suppose that:

- Recipient A is assigned by the Proposer the R-200 role. Recipient B is assigned by the Proposer the R-100 role.
- The Proposer offers Recipient A 80 francs (out of 200)
- The Proposer offers Recipient B 40 francs (out of 100)
- Recipient A *chooses a Minimum Acceptable Offer of 60 francs*
- Recipient B *chooses a Minimum Acceptable Offer of 50 francs*

This example results in the following payoffs:

- Recipient A:

In this case, the Proposer offered 80 francs, which is more than 60 francs, the minimum amount Recipient A would accept.

Payoffs: The Proposer receives 200-80=120 francs and Recipient A receives 80 francs.

- *Recipient B:*

In this case, the Proposer offered 40 francs, which is less than 50 francs, the minimum amount Recipient B would accept.

Payoffs: The Proposer receives 0 francs and Recipient B receives 0 francs.

- *Proposer:*

From above, the Proposer receives 120 francs from his/her interaction with Recipient A(the remainder of 200 francs), and receives 0 francs from his/her interaction with Recipient B(as the Proposer's offer was rejected). Thus in total the Proposer receives 0+120=120 francs.

Summary

If you are randomly selected to be the Proposer you will first have to assign the R-200 role and the R-100 role to the Recipients. Then you will have to choose an offer for each of the two Recipients. If you are randomly selected to be a Recipient, you will learn about the other Recipient's offer and will then have to state the minimum offer you are willing to accept.

Are there any questions?

Appendix B. The Quiz

1. Who was the first man in space?

   A. Neil Armstrong
   B. **Yuri Gagarin**
   C. Alan Shepard
   D. Dennis Tito

2. If the symbol of your profession is a “caduceus”, you are a:

   A. Dentist
   B. Detective
   C. **Doctor**
   D. Dog trainer

3. When did New Zealand move from pounds and shillings to dollars and cents

   A. **1967**
   B. 1961
   C. 1971
   D. 1973

4. What message did Benito Mussolini famously send to the Italian soccer team before the 1938 World Cup?

   A. We shall overcome
   B. You've come, you've seen, now conquer!
   C. To victory, for the blood of your forefathers
   D. **Win, or die**

5. Which chemist, engineer, and armaments manufacturer invented dynamite and a detonator for it and then founded a longstanding series of world prizes for, among other things, promoting peace?

   A. Alfred North Whitehead
   B. Alfred Russel Wallace
   C. Alfred, Lord Tennyson
   D. **Alfred Nobel**

6. Which of these is a prefix that means "false"?

   A. **Pseudo**
   B. Rhino
   C. Midi
   D. Mega

7. Who lived in the Lateran Palace from the fourth to the fourteenth centuries?

   A. Bavarian Kings
   B. Roman emperors
   C. **Popes**
   D. Prisoners

8. Which of these is not a prime number?

   A. 59
   B. **69**
   C. 79
   D. 89

9. These are the first lines from what song: "The falling leaves drift by my window, The falling leaves of red and gold, I see your lips, the summer kisses, The sunburned hands I used to hold"?

   A. The Girls of Summer
   B. California
   C. Monday Monday
   D. **Autumn Leaves**

10. How are coffee beans harvested?

    A. **Plucked from a tree like apples**
    B. Dug up from the ground like potatoes
    C. Threshed from stalks like wheat
    D. Picked up from the ground like cabbages

11. What was the largest cruise liner ever constructed when it was completed in 2003?

    A. Queen Victoria
    B. Queen Elizabeth II
    C. Queen Anne II
    D. **Queen Mary II**

12. In the Godfather, what is the surname of the characters Vito and Sonny?

    A. Nobilio
    B. Neri
    C. Pasquloni
    D. **Corleone**

13. What is a balalaika?

A. **An instrument**
B. A shellfish
C. A garment
D. A dish

14. What is Williams and Kate's newborn royal baby called?

A. **George Alexander Louis**
B. George Louis David
C. George Henry James
D. George Albert James

15. Which company has bought Nokia's mobile phone business?

A. Apple
B. Vodafone
C. **Microsoft**
D. Verizon

16. Who was the Greek goddess of the hunt?

A. Minerva
B. **Artemis**
C. Hera
D. Venus

17. In what year was the United Nations established?

A. 1947
B. **1945**
C. 1941
D. 1950

18. In which city would you find the Taronga Zoo?

A. Chicago
B. **Sydney**
C. Taronga
D. London

19. What is the name of Simba's mother in "the Lion King"?

A. **Sarabi**
B. Nala
C. Sarafina
D. Nobami

20. Arthur Guinness is famous for founding what?

A. The Guiness World Records
B. **Guiness Brewery**
C. Guiness Cake
D. Guiness Park

Appendix C. The Questionnaire

Please state your gender.

How old are you?

What is your nationality?

How many siblings do you have?

If you are a student, what is your subject?

When you were 16 years of age, what was the income of your parents in comparison to other families in New Zealand?

How large was the community where you have lived the most time of your life?

How many people live in your household (including yourself)?

How large is your monthly budget (without expenses for accommodation)?

What share of your monthly expenses do you finance yourself?

Can you state the percentage we can rely on the data you provided?

Did you decide on the wage offer based on the quiz score?

(only Principals, Treatment T1 and T3)

Do you think the principal's decision was fair?

(Recipient A, Recipient B, Treatment T1, T2 and T3)

Do you think the principal assigned roles based on quiz scores of the two agents?

(Recipient A, Recipient B, Treatment T1 and T3)

Do you find an allocation using the test scores as a fair way to make allocation?

Did you assign the low productivity role and the high productivity role based on the quiz score?

(only Principals, T2)

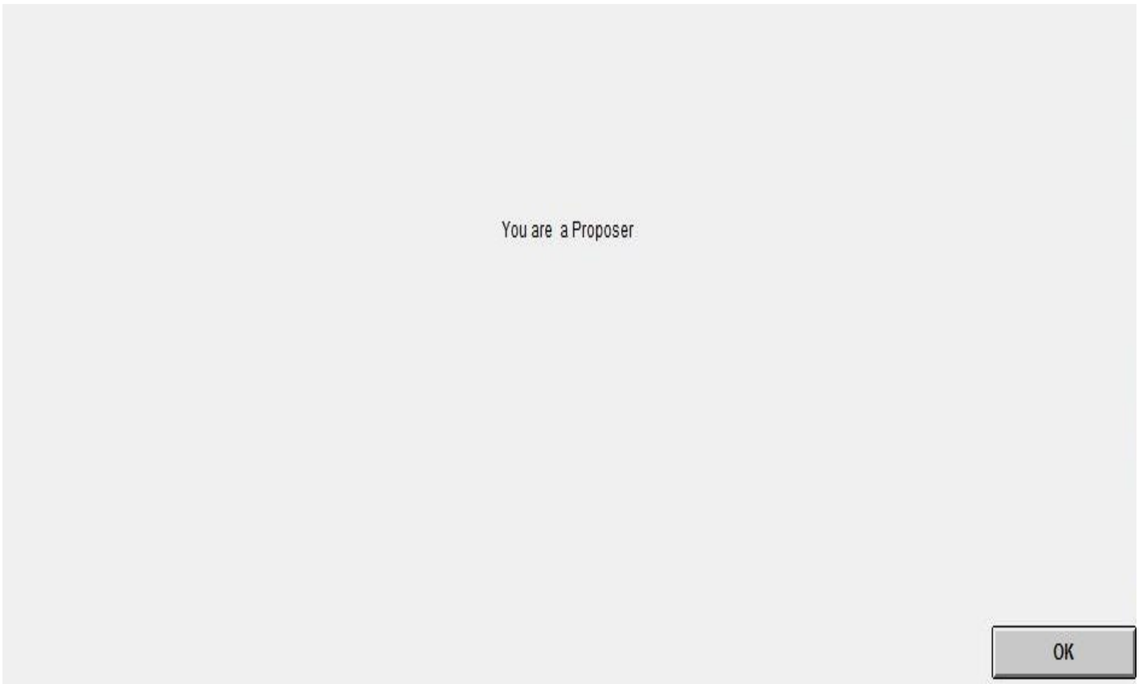
Do you think the principal assigned roles based on quiz scores of the two agents?

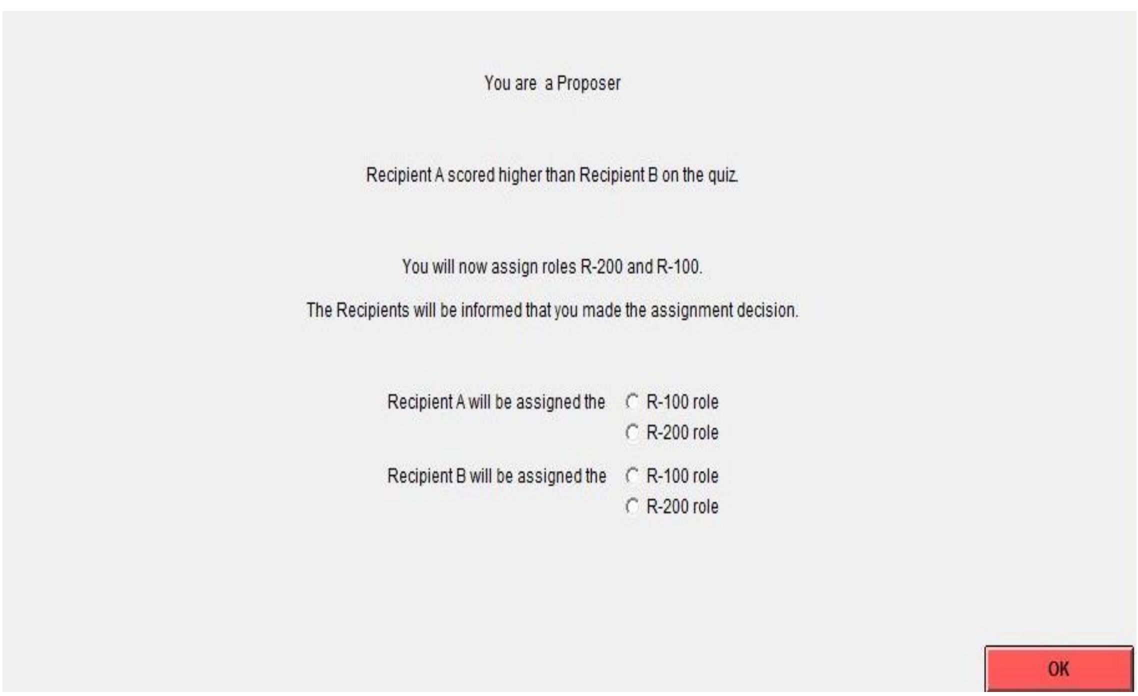
(Recipient A, Recipient B, Treatment T2)

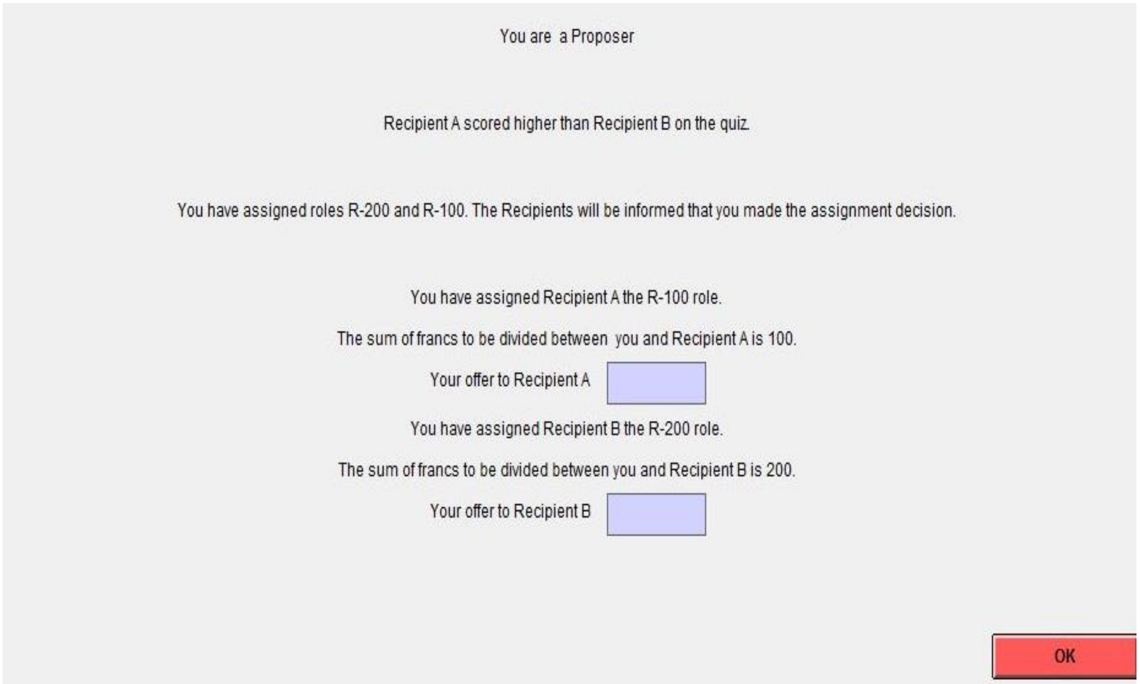
Appendix D. Screenshots From the Software

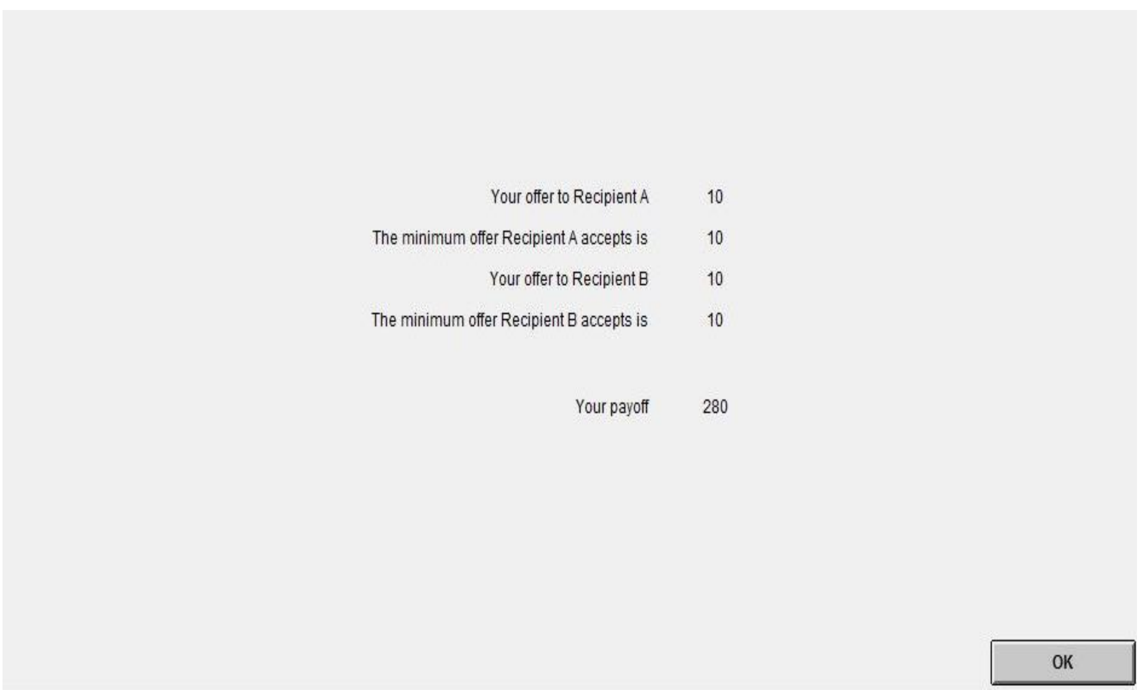
The Principal Assignment Treatment

The Principal (called The Proposer in the experiment)

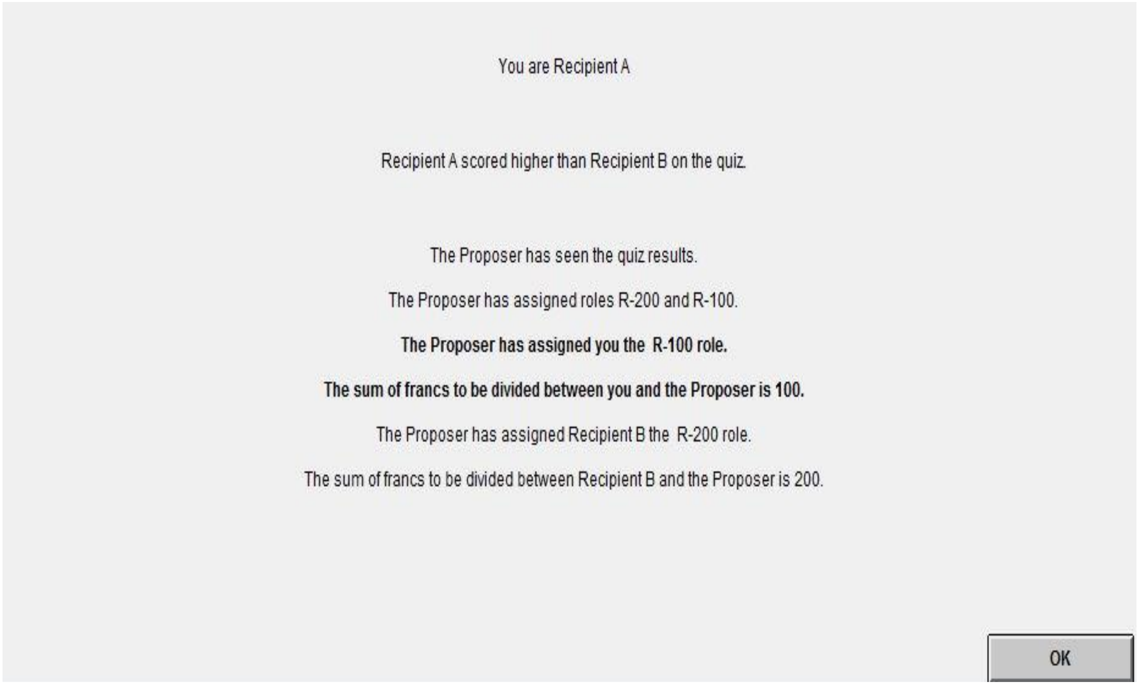


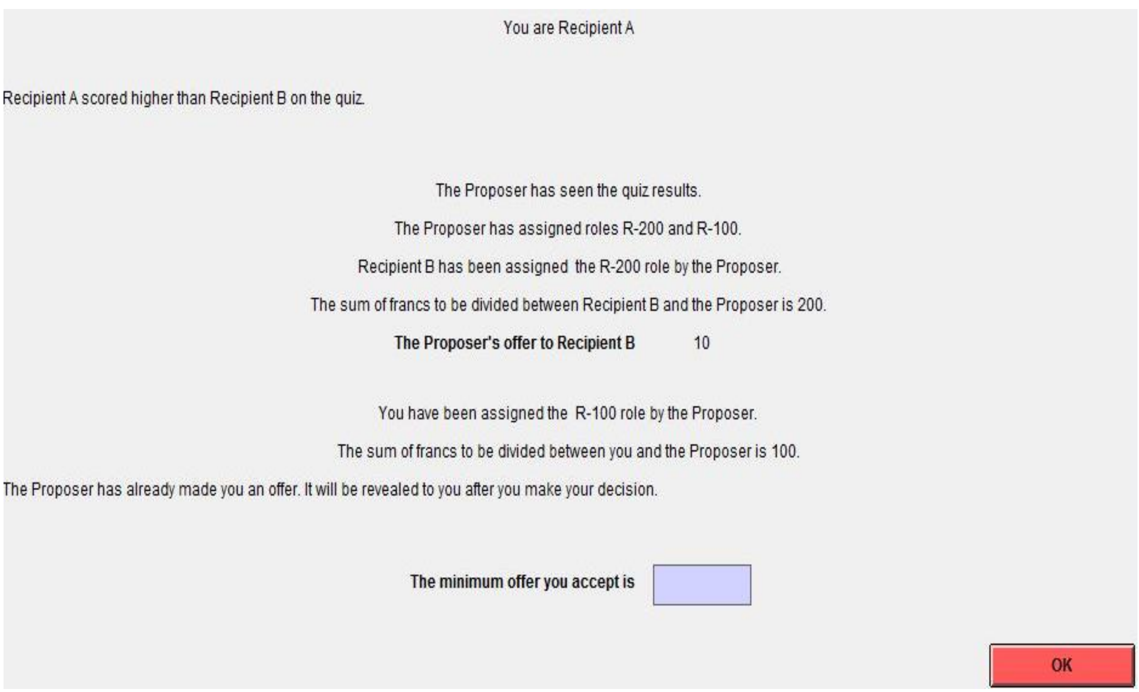


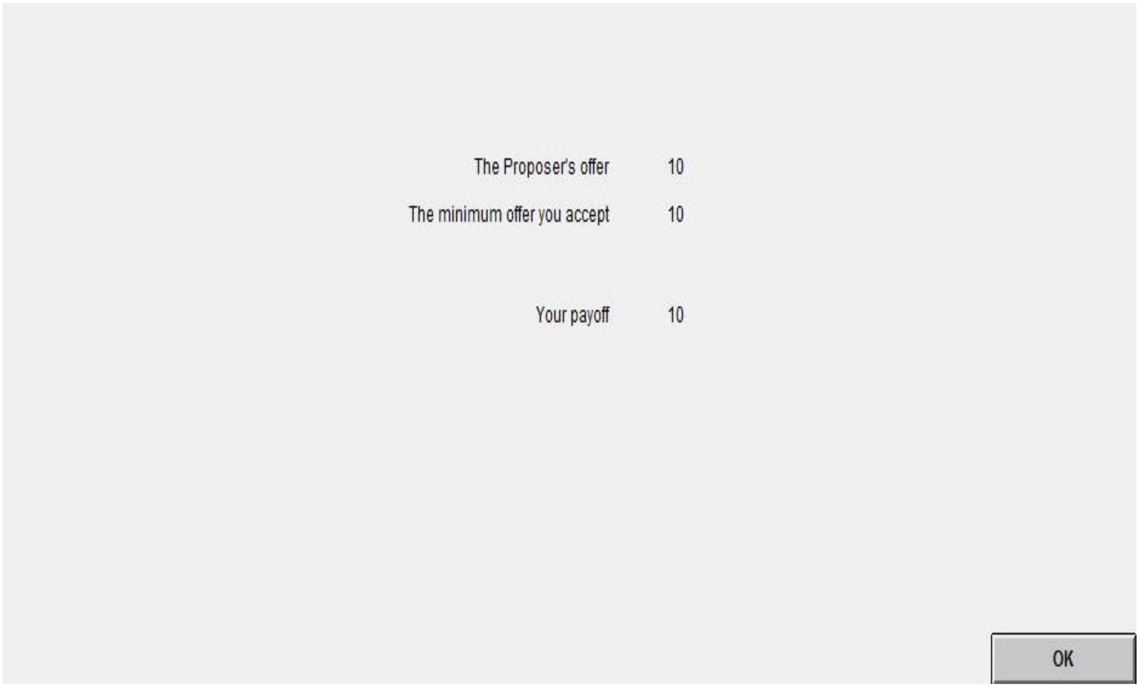

You are a Proposer
Recipient A scored higher than Recipient B on the quiz.
You have assigned roles R-200 and R-100. The Recipients will be informed that you made the assignment decision.
You have assigned Recipient A the R-100 role.
The sum of francs to be divided between you and Recipient A is 100.
Your offer to Recipient A
You have assigned Recipient B the R-200 role.
The sum of francs to be divided between you and Recipient B is 200.
Your offer to Recipient B
OK


Your offer to Recipient A 10
The minimum offer Recipient A accepts is 10
Your offer to Recipient B 10
The minimum offer Recipient B accepts is 10
Your payoff 280
OK

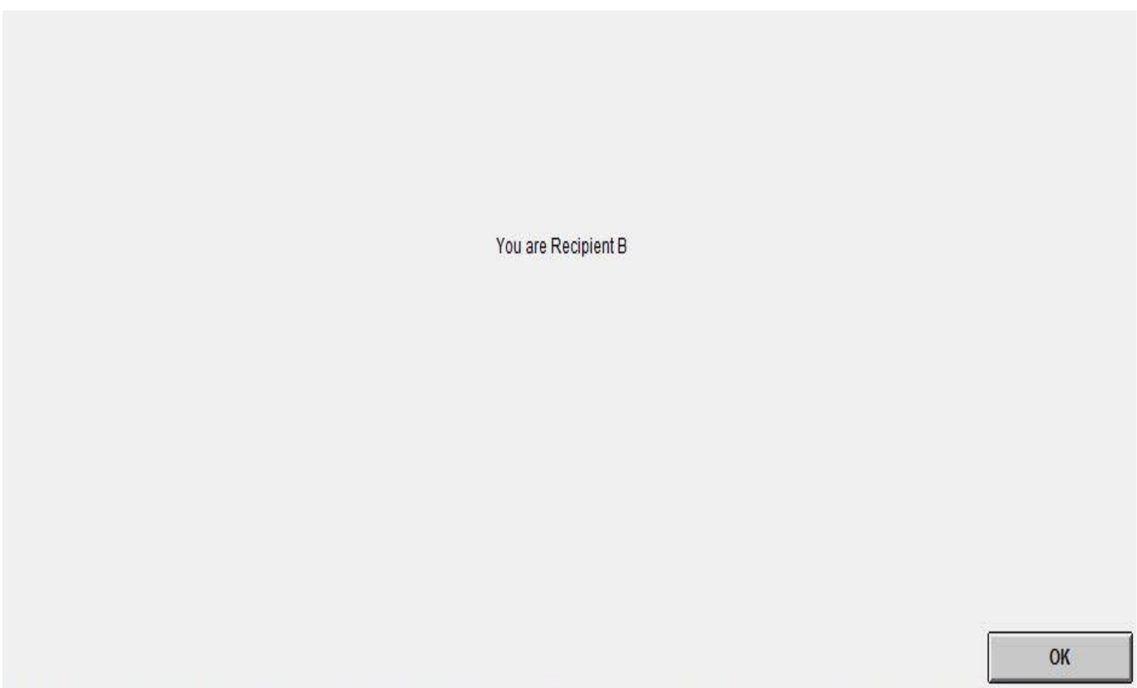
Recipient A:

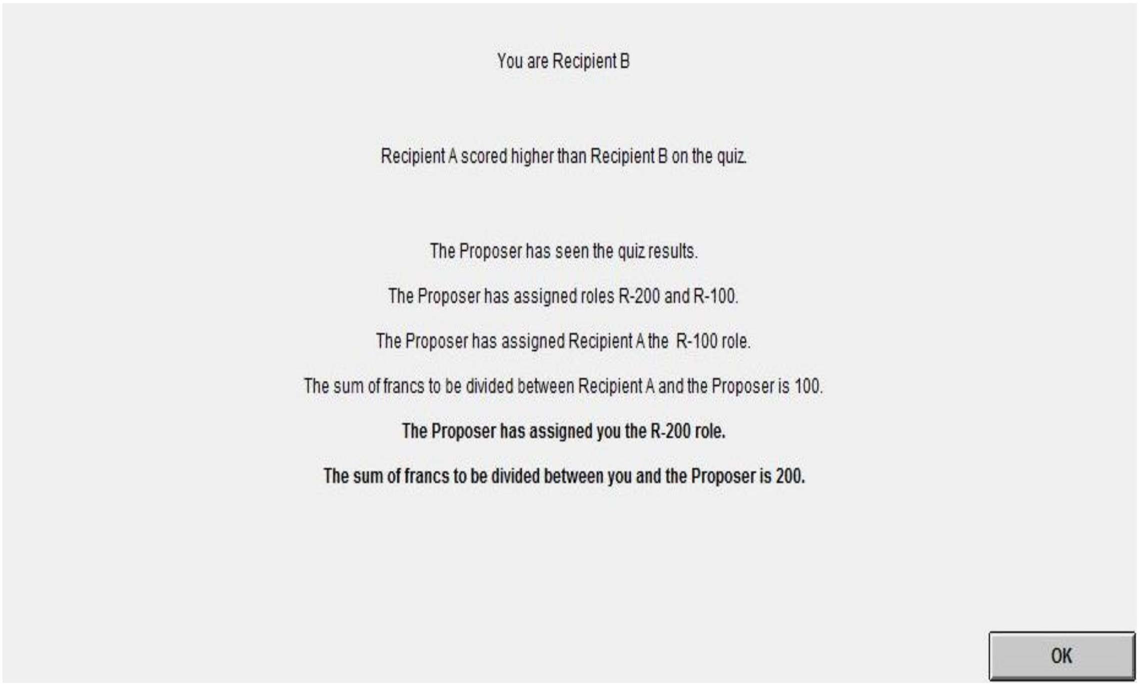

You are Recipient A
Recipient A scored higher than Recipient B on the quiz.
The Proposer has seen the quiz results.
The Proposer has assigned roles R-200 and R-100.
The Proposer has assigned you the R-100 role.
The sum of francs to be divided between you and the Proposer is 100.
The Proposer has assigned Recipient B the R-200 role.
The sum of francs to be divided between Recipient B and the Proposer is 200.
OK

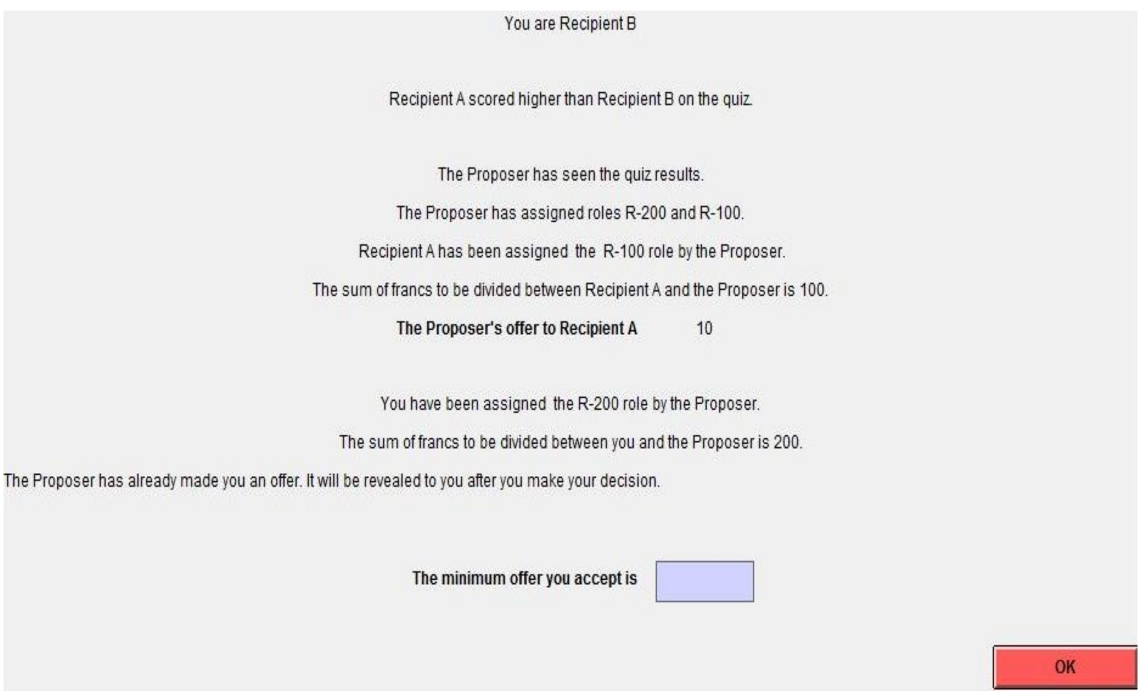

You are Recipient A
Recipient A scored higher than Recipient B on the quiz.
The Proposer has seen the quiz results.
The Proposer has assigned roles R-200 and R-100.
Recipient B has been assigned the R-200 role by the Proposer.
The sum of francs to be divided between Recipient B and the Proposer is 200.
The Proposer's offer to Recipient B 10
You have been assigned the R-100 role by the Proposer.
The sum of francs to be divided between you and the Proposer is 100.
The Proposer has already made you an offer. It will be revealed to you after you make your decision.
The minimum offer you accept is
OK

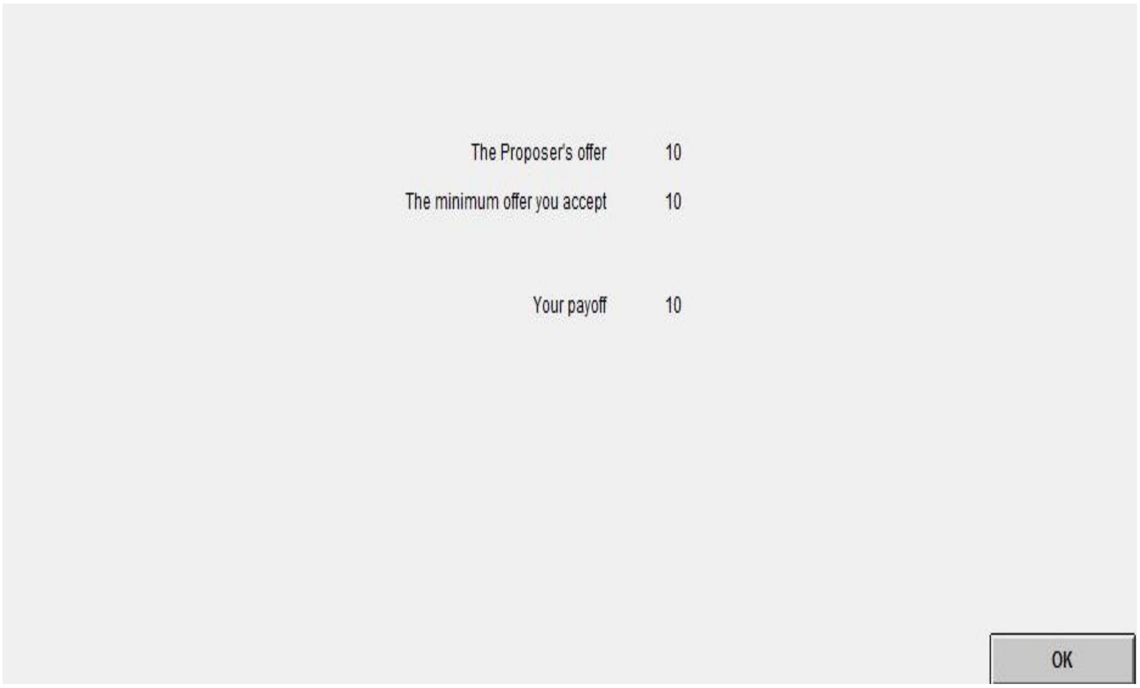

The Proposer's offer 10
The minimum offer you accept 10
Your payoff 10
OK

Recipient B:


You are Recipient B
OK


You are Recipient B
Recipient A scored higher than Recipient B on the quiz.
The Proposer has seen the quiz results.
The Proposer has assigned roles R-200 and R-100.
The Proposer has assigned Recipient A the R-100 role.
The sum of francs to be divided between Recipient A and the Proposer is 100.
The Proposer has assigned you the R-200 role.
The sum of francs to be divided between you and the Proposer is 200.
OK


You are Recipient B
Recipient A scored higher than Recipient B on the quiz.
The Proposer has seen the quiz results.
The Proposer has assigned roles R-200 and R-100.
Recipient A has been assigned the R-100 role by the Proposer.
The sum of francs to be divided between Recipient A and the Proposer is 100.
The Proposer's offer to Recipient A 10
You have been assigned the R-200 role by the Proposer.
The sum of francs to be divided between you and the Proposer is 200.
The Proposer has already made you an offer. It will be revealed to you after you make your decision.
The minimum offer you accept is
OK


The Proposer's offer 10
The minimum offer you accept 10
Your payoff 10
OK

Appendix E. Human Ethics Committee Approval

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Ref: HEC 2014/35

13 June 2014

Katarina Dankova
Department of Economics & Finance
UNIVERSITY OF CANTERBURY

Dear Katarina

The Human Ethics Committee advises that your research proposal "Job assignment and wage comparison" has been considered and approved.

Please note that this approval is subject to the incorporation of the amendments you have provided in your email of 3 June 2014.

Best wishes for your project.

Yours sincerely

Lindsey MacDonald
Chair
University of Canterbury Human Ethics Committee

Chapter 3

3 Overconfidence and Excess Entry

3.1 Introduction

This chapter seeks to replicate a study by Camerer and Lovallo (1999) (henceforth CL) on “Overconfidence and Excess Entry”. CL explore overconfidence from a behavioural economics point of view. They test whether optimistic biases could predictably influence economic behaviour of firms when entering into markets. Widespread overconfidence would have important consequences on economic behaviour. Overconfidence in one's skills or relative ability can, in financial markets, lead to excessive trading and lower returns (Barber and Odean, 2001), distortions in corporate investment decisions (Malmendier and Tate, 2005), value-reducing mergers (Roll, 1986) and to various security market anomalies (Daniel, Hirshleifer, and Subrahmanyam, 1998).

Overconfidence of entrepreneurs is possibly crucial for understanding business failures. In general, the business failure rate varies according to industry. The US Bureau of Labor Statistics reports the business survival rate of companies established in 2006 was 49.3% in 2011.[24] The survival rate decreases with the number of years of the company existence. For example, only 24.6% of companies established in 1994 were still in the market in 2011. CL offer three possible explanations for entrant failure. The first one states that failures are hit-and-run entries that are profitable but brief. Profits are

[24] The U.S. Bureau of Labor Statistics collects data on the nation’s labour market, i.e. new businesses and job creation, Business Employment Dynamics, http://www.bls.gov/bdm/entrepreneurship/bdm_chart3.htm.

made if entering the market during the high peak, i.e. when profitability is high and then leaving (or 'failing') when profitability reduces. Because of the fleeting nature of many business opportunities, a failure within a year of setup is probable and expected (Forbes, 2009). In the second explanation CL compare business entries to lottery tickets, i.e. most firms expect to lose money and fail, but if they become successful, the payoff is very large and worth the risk. The third explanation states that many entry decisions are simply mistakes, due to either underestimating the competitors or overconfidence about own abilities. It can be challenging to distinguish which one of the three mentioned explanations influences business failure and to what extent. CL designed an economics experiment testing for the effect of overconfidence in one's skill on market entry decisions, i.e. if overconfidence amplifies the entry rate.

Menkhoff, Schmidt, and Brozynski (2006) used a survey to investigate overconfidence. They conducted a questionnaire survey regarding the impact of experience on risk taking, in particular whether inexperienced fund managers tend to take higher risks than their experienced colleagues. Higher risk taking may be explained by a higher degree of overconfidence, a lower degree of risk aversion, or less herding behaviour.[25] They find that herding decreases with experience, whereas overconfidence decreases with experience for some tasks and increases for others. Also, Ben-David, Graham, and Harvey (2007), based on their survey, tested whether top corporate executives are overconfident (their measure of overconfidence is based on miscalibration of beliefs) and whether their overconfidence impacts investment behaviour.

[25] Herding can be defined as the phenomenon of individuals deciding to follow the choices others and imitating group behaviour rather than deciding independently on the basis of private information (Baddeley, 2010).

To test for overconfidence in market entry CL employ experimental economics methods, which by motivating people financially result in salient choices rather than non-salient answers and thus offer better control over the data generating process than surveys. Economics experiments differ from psychology experiments and surveys in that Economics experiments pay performance-based incentives, which reduces response noise, e.g. extreme outliers probably caused by thoughtless, unmotivated subjects (see Smith and Walker, 1993). Economists presume that experimental subjects do not work for free. They work harder, more persistently, and more effectively, if they can earn more money for better performance. On the other hand, psychologists believe that intrinsic motivation is usually high enough to produce steady effort even in the absence of financial rewards; and while more money might induce more effort, the effort does not always improve performance, especially if good performance requires subjects to induce spontaneously a principle of rational choice or judgment, like Bayes' theorem, where subjects update their probabilities (Camerer and Hogarth, 1999). In addition, Economics experiments prohibit deceiving subjects. In experiments, the experimenter has to be able to control what the subjects' expectations are when testing theoretical equilibria. For example, there is a 1 in 6 chance of rolling any number using a fair die and this is a common knowledge. With deception control of expectations is weakened. Avoiding deception is often believed to be a hallmark of Experimental Economics (Davis and Holt, 1993) and is duly enforced by editors and reviewers (Roth, 2001). All deception in Experimental Economics is forbidden (Wilson and Isaac, 2007) and experiments with deception are normally not publishable in any economics journals (Gächter, 2009). In addition, it is widely believed that such a ban is to avoid reputational spillover effects, meaning that deception might compromise the reputation of the experimental laboratory if the participants expect to be deceived (Hertwig and Ortmann, 2001). There is no deception in my experiment, i.e. the subjects receive true information about the payoffs, the market capacity, their ranks and are not deceived at any point during the experiment.

The idea that overconfidence causes business entry mistakes has been suggested before (e.g. Roll, 1986), but CL were first to directly test it by measuring economic decisions and personal overconfidence at the same time. CL in their study extend the market entry game described in Kahneman (1988), in which subjects decide simultaneously and without communication whether they want to enter the market or not.[26] In Kahneman's market entry game, the payoffs depend on the pre-announced market capacity and the number of entrants.[27] Only if the number of entrants is lower or equal to the market capacity, entrants are making money. After a few rounds, the number of entrants in each round approximates to the market capacity even if entrants are not communicating. In order to test for overconfidence, CL add more features to Kahneman's market entry experiment. In CL entrant's payoff depends on the market capacity, the number or entrants as well as on an entrant's rank relative to other entrants. The ranks are assigned both randomly and according to the entrant's skill, which is determined by his relative performance on a quiz. The only decision the subjects make is whether they wish to enter a market or not.

I ask the two same main questions as CL: 1. Is there more entry when people are betting on their own relative skill? 2. Are subjects neglecting the reference group, when

[26] The market entry game, originally called N*game, was invented by Kahneman, Brander and Thaler. This game was never officially published, it is only informally described in Kahneman (1988). Since then it has been used by others, e.g. Erev and Rapoport (1998).
[27] The payoffs in Kahneman's experiment are calculated according to a following formula: $0.25(capacity - number of entrants).

they volunteer to participate in the experiment knowing that their payoffs will depend on their skill? The answers to these questions can help us obtain the performance feedback necessary to correct optimistic overentry.

CL find that when payoffs depend on skills, overconfidence and excess entry are higher than in situations when payoffs are determined randomly. Furthermore, excess entry is highest when the subjects were told in advance that their payoffs will depend on their skill, suggesting that subjects were neglecting consideration of the reference group with which they would be competing.[28] In my experiment I obtained the same results as CL only to some extent using different and more suitable econometric tests. I have implemented CL's experiment, with some modifications reflecting the technological progress of Economic experiments over the past 15 years. My experiment is computerized as opposed to CL's pen and paper. I also made a change to the subject pool; my experiment studies the effect of overconfidence of both genders, making it a more conservative test. CL used exclusively male subjects in some sessions in order to control for gender confound as are usually less overconfident than men. I have, however, preserved the same order of sessions with the same number of participants and same parameters.

[28] This is also known in the literature as egocentrism (Kruger, 1999; Windschitl and Chambers, 2004).

3.2 Replications In Experimental Economics

The seventh commandment of Croson and Gächter (2010) "ten commandments" of economic science states: *"Experimentalists shall replicate and encourage replications, including making their data, instructions and software publically available."*

What does replication in the context of economic research mean? Hunter (2001) distinguishes between pure and scientific replication. Pure replication is based on running the test using the same data as previous papers. Pure replication depends on the availability of all the information from the project that is to be replicated. Making pure replications possible allows important findings to be verified or refuted directly using the particular data set used to generate them, rather than examined later on a different and perhaps less appropriate set of data. Technology has diminished the costs of providing the materials necessary for pure replication. Most of what economists view as replication represents scientific replication, i.e. re-examining an idea in some published research by studying it using a different data set chosen from a different population from that used in the original paper (Hamermesh, 2007). The purpose of replications is proving that an experimental phenomenon is robust to small or large changes in the experimental setup. A result is robust when it does not depend on some detail of the situation or on the assumptions used to derive it. A "good" scientific result is always robust to some kind of variation or change (Guala, 2005). Kuhn (1962) noticed that for long periods of time scientific research was characterized by the accumulation of results that contribute to the growth of knowledge without challenging the basic tenets of the received view. In Experimental Economics, replication is the capacity to create an entirely new set of observations. The availability of detailed design description, thorough procedures and software is important in order gain insights into data generating process, which is to be preserved when replicating a study. Public goods experiments are among the most widely replicated in economics, experimental psychology, sociology and political science (Guala, 2005).

In natural sciences replication is a crucial activity and contributes to quality of the research. In economics, even though replication of data has been discussed thoroughly, publication of replication studies has been slow. In many cases replications do not confirm the results of the original studies. The journals might thus be afraid of the disagreements between replicator and the original authors. Ragnar Frisch in the first issue of *Econometrica* in 1933 noted the importance of making data available to researchers in order to enable replication. Substantial progress with data sharing and replications has been made over recent years. Journals such as the *American Economic Review*, *Econometrica*, *Experimental Economics* and the *Journal of Applied Econometrics* now have official data sharing and replication policies in effect. Similarly, *Public Finance Review* and the new *Journal of the Economics Science Association* promote economic replications. In these as well as in an increasing number of other journals, authors are expected to state in which public archive they will deposit the information necessary to replicate their numerical results. This archive should include original data, specialized computer programs, extracts of existing data files and explanatory files describing how to reproduce the exact numerical results in the published work. The changes in the publication process in economics have increased the benefits to authors of maintaining the records that might make replication possible. Replications in economics are required and promoted, however, they are still quite rare because the incentives are low (Dewald, Thursby, and Anderson, 1986). The lack of publishing outlets is perhaps the most serious obstacle to researchers interested in undertaking replication research (Duvendack, Palmer-Jones, and Reed, 2015).

It is important to mention the difference between a replication and a robustness check as these terms are often blurred. In econometrics, a replication test estimates parameters drawn from the same sampling distribution as those in the original study (Clemens, 2015). A replication test can take two forms: a verification test and a reproduction test. A verification test ensures that the exact statistical analysis using the original dataset. A reproduction test resamples the same population but otherwise using identical methods to the original study (Clemens, 2015). A robustness test, as opposed to replication, estimates parameters drawn from a different sampling distribution from those in the original study. A robustness test can take two forms: a reanalysis and an extension. A reanalysis test alters the computer code from the original study. An extension test uses new data gathered from a different population sample, or gathered on the same sample at a substantially different time.

Replication is Experimental Economics is different to the one in Econometrics. When one group tries to replicate an experimental result reported by another, it does not usually replicate every detail of the original design. Instead, it adapts that design to its own methods of recruitment, instructions, display format, methods of payments, etc. (Bardsley, 2010). Reproducibility of the experimental results is related to the control over the relevant variables. Replicating an experiment can serve to test the robustness of past results or to test the effects of additional parameters. The fact that the results are easily replicable contributes to the expansion of the Experimental Economics field (Roth, 1994). When a carefully conducted experiment is repeated the likelihood of data being similar is high. Roth (1994) distinguishes between precise replication and a replication with variation of experimental parameters and conditions. The second gives us best indication of the robustness of the experimental results.

To my knowledge the paper by CL has not been replicated previously. For the purposes of my thesis I have chosen to replicate their experiment using computers, and a mixed-gender subject pool as there is little evidence that gender influences trading activity (Deaves, Luders, and Luo, 2009). I find this approach contributes to the literature by adding an example of comparing a pen and paper experiment to a computerized one. Also, programming what was originally a pen and paper experiment and making the software available, provides accountability and ease of replication.

3.3 Literature Review

Overconfidence has been studied in psychology for many years, though less so in economics. Yet in the past, microeconomics was closely linked to psychology. For example, Adam Smith, who is considered as being one of the pioneers of economics, in his book *The Theory of Moral Sentiments*, mentions that we suffer more when we fall from a better to a worse situation than we enjoy when we rise from worse to a better. Even though the exact term is not stated in the book, this is what we know today as loss aversion. In this book he proposed psychological explanations of individual behaviour, including concerns about fairness and justice. There is a long history of economists seeking alternatives partially in psychology because the classic economic assumptions of unbounded rationality, unbounded willpower, and unbounded selfishness have been violated (e.g. the effects of cognitive biases on markets in Arrow, 1982; and Camerer, 1992; fairness and the acceptability of the business transactions in Kahneman, Knetsch, and Thaler, 1986; differences between willingness-to-pay and willingness-to-accept in Kahneman, Knetsch, and Thaler, 1990; violations of description invariance in Tversky and Kahneman, 1986). Behavioural economics increases the explanatory power of economics by providing it with more realistic psychological foundations (Camerer, Loewenstein, and Prelec, 2004).

Overconfidence occurs when an individual's certainty that his predictions are correct exceeds the accuracy of those predictions (Klayman et al., 1999). Participants in psychology surveys often report that they expect to have higher starting salaries, job satisfaction, more enduring marriages, more gifted children and other positive life events compared to the average population (Hoorens and Buunk, 1993; Perloff and Fetzer, 1986;

Weinstein, 1980). Conversely, participants report a below-average risk of being robbed or assaulted, or experiencing unemployment, job loss or health problems (Miller et al., 1990; Weinstein, 1980; Weinstein, 1982; Weinstein, 1984; Weinstein, 1987; Weinstein, 1989; Weinstein, 1998; Weinstein and Klein, 1995). Hoelzl and Rustichini (2005) state that people may be overconfident in many different ways: they may overestimate their abilities, perceive themselves more favourably than others perceive them, or finally, perceive themselves more favourably than they perceive others. There is a large body of evidence in psychology and social psychology showing that people are overconfident about their relative abilities or unreasonably optimistic about their future (Alicke, 1985; Dunning, Meyerowitz, and Holzberg, 1989; Messick et al., 1985; Taylor and Brown, 1988). The effect has been labeled "better than average". A popular example of overconfidence is asking a group of average people about their driving ability. Most of them will say they are above average even though only about half can be better than average (Svenson, 1981).

The literature in economics and finance on theories, implications and practical effects of overconfidence is growing to be large and influential as it can explain economic behaviour.[29] For example, De Long, Summers, and Waldmann (1991) show that overconfident investors, by taking more risky positions with higher returns, may come to dominate asset markets. The economic literature has emerged from psychological studies over the past thirty years. Odean in his theory (Odean, 1998; Odean, 1999) states that investors are too confident about their own ability and will trade too much because they give too much credence to their own private signal about value,

[29] For instance, see Bénabou and Tirole (2002); Bénabou and Tirole (2003); Daniel, Hirshleifer, and Subrahmanyam (1998); Gervais and Odean (2001); Weinberg (2009).

while failing to recognize other investors' private signals. He tested this hypothesis on 10,000 clients of a major US brokerage service from 1987 to 1993 and found that the investors do indeed trade too much (Odean, 1999). De Bondt and Thaler (1995), in their detailed summary of the micro foundation of behavioural finance, express that the finding that people are overconfident might be one of the most robust in the psychology of judgement. One of the first experiments testing overconfidence in the market entry was designed and ran by CL. They find that when payoffs depend on skill, more subjects enter the market (thus are more overconfident) than in a situation when payoffs are determined randomly. Furthermore, excess entry is highest in sessions for which it was common knowledge that subjects' payoff will depend on their skill, suggesting that subjects were neglecting consideration of the reference group with which they would be competing.

Theoretical models distinguish between optimistic managers who overestimate the mean of their firms' cashflows (Hackbarth, 2008; Heaton, 2002; Shefrin, 2001) and overconfident managers who either underestimate the volatility of their firm's future cash flows or overweight their private signals relative to public information (Gervais and Goldstein, 2007; Gervais, Heaton, and Odean, 2011). Heaton (2002) in his model on an optimism bias, that might be affecting the decisions of managers, explains two trends that are often observed in investment policy. First, optimistic managers believe that financial markets undervalue their firm and second, managers overestimate their own ability to manage new projects. Barber and Odean (2000) document that losses are attributable to overconfidence, finding that those trading the most earn the least on a risk-adjusted basis.

Overconfidence is also studied with regard to a firm's optimal wage policy. CL find that subjects are more overconfident when betting on their own skill rather than relying on a random device. How does this fact influence workers' behaviour in a firm? Fang and Moscarini (2005), for example, provide a theoretical link between workers' overconfidence and wage-setting practices of the firm. The firm can either offer a wage based on a worker's performance (differentiation policy), or offer the same wage to all workers (non-differentiation policy). Workers' confidence in their own skills is interpreted as their morale, i.e. a worker has a high morale when he believes that his effort is meaningful and has impact on output of the firm. On the other hand, a worker is demoralized when he believes that his costly effort is useless. In their model, wages provide incentives and affect worker morale, by revealing private information of the firm about worker skills. Their model provides conditions for the non-differentiation wage policy to be profit maximizing. In many examples, worker overconfidence is a necessary condition for the firm to prefer the non-differentiation wage policy, so as to preserve some workers' morale. The non-differentiation wage policy itself triggers more workers' overconfidence, i.e. stronger belief that one's effort is meaningful. This result thus serves as a base for firms' wage secrecy as documented by Bewley (1999).

Apart from theoretical evidence of the importance of overconfidence, there are also numerous economics experiments testing for the presence of overconfidence. The closest to CL is the study by Simon and Houghton (2003). They conducted a field study examining the role of overconfidence in actual entry decisions, such as product introductions in the computer industry. They test whether managers introducing pioneering (i.e. risky) products express extreme certainty more or less frequently than managers introducing fewer pioneering products. Using the Georgia Technology

Sourcebook they identified 135 small computer companies that anticipated launching a new product within 30 days or had just launched a product within the past three months. They interviewed the managers, who were asked to describe what factors they believed were important in order for their product introduction to succeed. The questions about the production introduction decision were disaggregated into discrete components rather than asking about the overall success of the product introduction. They first collected data from each participating firm around the time the company launched its product to assess the extent to which the product was pioneering, the success factors the manager was focusing on, and the manager's level of certainty in achieving each success factor. Eighteen months later they collected data to determine whether the new product introduction had achieved the specified success factors. Pioneering products are riskier than incremental product introductions, because pioneering introductions are unique, lacking prior similar actions to help calibrate judgment (Golder and Tellis, 1993). Managers use a diagnostic cue when attempting to forecast the success of their strategic actions.[30] A diagnostic cue is an indicator that is present the majority of the time given one outcome and absent the majority of the time given the alternative outcome (Soll, 1996). A diagnostic cue may be a stronger predictor of success in common decision contexts but weaker in uncommon decision contexts. Overconfidence occurs when managers overestimate the predictive validity of the cue. On the one hand, managers in pioneering decision contexts might be less likely to express extreme certainty than managers in other decision contexts, because diagnostic cues are weaker predictors of success in pioneering contexts. On the other hand, managers in pioneering decision contexts may be relying on a few salient examples of past situations when the diagnostic

[30] A diagnostic cue can be, for example, positive customer feedback prior to the introduction of a product.

cue was associated with success in order to determine their level of certainty about their own prediction. Subsequently, they found that overconfidence is associated with introducing products that are more pioneering than incremental. They also found that managers who were extremely certain that they would achieve success were more likely to introduce pioneering rather than incremental products. Yet in outcomes, success was associated less with pioneering product introductions than with incremental introductions, which provides evidence of overconfidence similarly as in CL where subjects are overconfident when betting on their own skill rather than on a random device.

The link between overconfidence and trading activity on an individual level was experimentally tested by Deaves, Luders, and Luo (2009). They have three variants of their computerized double auction market: the calibration based variant, the better than average effect variant and the illusion of control variant. Deaves, Luders, and Luo (2009) use a pre-experimental questionnaire including general knowledge questions, which are designed to influence behaviour and serve as a measure for calibration-based overconfidence. In addition to obtaining demographics subjects had to answer twenty general knowledge questions, which had objectively known numerical answers. The main purpose was to find out how well calibrated subjects were. Afterwards, they were assessed both on their level of knowledge and on their degree of calibration-based overconfidence. Knowledge was measured by summing up absolute differences between midpoints of confidence intervals and correct answers (scaled by correct answers). Calibration-based overconfidence was assessed by calculating the percentage of times that confidence intervals contained correct answers. Deaves, Luders, and Luo (2009) find (similarly to Camerer and Lovallo (1999)) that greater overconfidence is likely to lead to aggressive behaviour in the pursuit of higher wealth. They also tested for gender effect as there are findings showing that men are more predisposed to overconfidence (Lundeberg, Fox, and Puncochar, 1994). The tendency to ascribe success to personal effort and failure to external forces is less pronounced from women (Beyer, 1990). They find little evidence that gender influences trading activity.

Although CL in their additional analysis outline the difference between overconfidence about one's skill and underestimating the performance of the other entrants, a detailed decomposition of overconfidence is provided by Hoelzl and Rustichini (2005). They recognize overconfidence resulting from three different types of behaviour: people overestimating their abilities; they perceiving themselves more favourably than others perceive them; or finally, they perceiving themselves more favourably than they perceive others. Since most of the evidence on overconfidence is based on verbal statements of subjective estimates, Hoelzl and Rustichini (2005) designed an experiment in which subjects have to choose among options, not statements. Subjects who are simply asked to evaluate their ranking have no incentive to be accurate and also might have an imprecise idea of concepts like percentile and median, or average.[31] The authors investigate how perceived relative skill influences verbal and choice behaviour. The revelation of the relative position that a subject thinks to have with respect to others is deduced from the choice itself. The study was a 2 x 2 between-subjects design. The treatment varied in payment, money vs. no money, and task difficulty, easy vs. difficult. In the money condition, participants could win 150 ATS.[32] In

[31] For example, Libby and Lipe (1992) studied recall and recognition of "controls", which accountants look for when auditing a firm. Subjects then had to recall as many of the controls as they could or recognize a control seen earlier. They found that incentives caused subjects to work harder and recall more items correctly.

[32] Austrian Schilling: 150 ATS were worth approximately 10 USD at the time.

the no money condition, they were asked to imagine that they could win the same amount. Before performing the task, subjects had to choose by a majority vote the procedure to determine their payment: either performance test or lottery condition. In the condition performance test a subject wins if his result were in the upper half of the results of all subjects. In the lottery condition he wins with 50% probability, with the outcome determined by an individual toss of a die. The vote determining the condition was confidential. Then the task of explaining easy or difficult words by completing sentences with gaps followed. What is relevant in subjects' choices is the estimated relative ability to solve the task. Hoelzl and Rustichini (2005) find that there is a sharp difference between the vote (measure of the degree of confidence) in the easy and difficult treatment. This difference is particularly strong when payment is offered. Choice behaviour changes from overconfidence to underconfidence when the task changes from easy to difficult.

Resulting from evidence on skill being important for overconfidence, Moore and Cain (2007) test whether the degree of difficulty of the task leads to overconfident choices. They show in their two experiments that people believe that they are below average on skill-based tasks that are difficult. Their design builds on that of CL's (1999) coordination game in which, in each round, N players decide whether to enter a market or not. Each entrant's payoffs depended on his or her rank within the market, such that more money is earned by better performance relative to other entrants. Each market had a pre-announced capacity, c. Entrants ranked below c lose money, entrants ranked c or above earned money, while non-entrants neither earned nor lost any money. The distinguishing feature of Moore and Cain's design is that skill-dependent payoffs are based on either an easy or a difficult trivia game. Contrary to the notion that overconfidence tends to be pervasive on all skill-based competitions, they predict that subjects will only believe they are better than others on simple tasks, on which they expect excess entry. They also test CL's explanation that people focus on themselves and simply neglect consideration of others (rather than miscalculating others' performance) when making comparative judgements. CL called this "reference group neglect". The reference group neglect explanation states that, for example, students neglect to consider the fact that other students are also likely to find the test difficult. The differential regression explanation used by Moore and Cain (2007), on the other hand, hypothesizes that people generally have better information about themselves than about others, so their beliefs about others' performances tend to be more regressive (thus less extreme) than their beliefs about their own performances. In other words, when your absolute performance is better (or worse) than your prior expectations, sensible Bayesian inference will lead you to make predictions of others' performances that are between these priors and your current beliefs about your performance. Reference group neglect refers to errors in the weight one puts on estimates of others' performance, while differential regression refers to errors in the estimate that is weighed. In the second experiment they test which interventions might be useful for reducing errors and which interventions are unlikely to be effective. Subjects were first told that they would be taking either a difficult or simple quiz and were then asked to predict the outcome. After taking the quiz, participants were invited to revise their answers to their prior estimates of absolute and relative scores. Finally, subjects were given full information about how others scored on the same quiz they took, and they were asked to report the same comparative judgements. Previous research has shown that information about others can reduce better than average effects (Alicke et al., 1995).

Accuracy of forecasting one's own performance, relative or absolute, was tested by Alicke, Klotz, Breitenbecher, Yurak, and Vredenburg (1995). In one task subjects were searching over a spreadsheet for the maximum value of a function, in the other subjects were decoding five-letter words. Participants were asked to predict how many rounds of a task contest they will win, an intuitive frequency-based method for eliciting a forecast of an unambiguous outcome. They find that underconfidence is more prevalent than overconfidence in forecasts of relative or absolute performance, which is in contrary to CL. CL find that subjects enter more often when betting on their own skill meaning that they expect to perform better than other subjects.

CL have triggered a new series of experiments testing overconfidence in different setups, and testing numerous causes leading to overconfidence, varying degrees of task difficulty that leads to overconfidence. The results from these experiments, for example, provide evidence that overconfidence is associated with introducing risky products to the market (Simon and Houghton, 2003), that greater overconfidence leads to aggressive behaviour in the pursuit of higher wealth (Deaves, Luders, and Luo, 2009) and that the difficulty of the task matters for overconfidence (Hoelzl and Rustichini, 2005; Moore and Cain, 2007). From economics point of view is inevitable that we understand how overconfidence influence choices and behaviour in order to avoid unnecessary business failures, to understand managers' decisions about investment policy and firms' optimal wage policies.

3.4 Experimental Design And Procedures

CL used for their design the market entry game described in Kahneman (1988). The game is played by a group of N people. In each round, the market capacity “c” is announced, where 0< c <15. The participants then make simultaneous choices of whether or not to enter the market. The payoff to the entrants depends on their number, according to the following formula: $0.25(c – number of entrants). Within very few trials, a pattern emerged in which the number of entrants, was within 1 or 2 of c, with no obvious systematic tendency to be higher or lower than c.

In CL as well as in my experiment 12 to 16 subjects (depending on a session) also decide in each of 24 rounds whether to enter a market or not. Each entrant starts with an initial endowment of $10. Each entrant's payoffs depend not only on the pre-announced market capacity “c”, but also on his rank. Entrants ranked below c lose their initial endowment, while entrants ranked c or above earn money (see Table 3.1). The top “c” entrants share $50 proportionally, with higher-ranking entrants earning more relative to other entrants.[33] Non-entrants do not earn or lose any money; they keep their initial endowment. If the number of entrants is exactly c+5, then the total payoff to all entering subjects is zero and if there are more than c+5 entrants, then on average entrants lose money. The ranks are assigned both randomly and according to the subject’s skill, which is determined by his relative performance in a quiz. Skill-ranks are determined by how fast the subjects finish five mazes (session 1 and 2) or how many trivia questions about sports or current events the subjects answer correctly (sessions 3-8). Subjects’ ranks are not revealed until the end of the experiment, i.e. after all have made their decisions.

[33] If the number of entrants is lower than c, the entrants share $50 proportionally, i.e. the entrant with the lowest rank receives the smallest $ amount, the entrant with the second lowest rank receives twice as much as the previous one etc.

Table 3.1 Rank-Based Payoffs*

	Market Capacity			
Rank	**2**	**4**	**6**	**8**
1	33	20	14	11
2	17	15	12	10
3		10	10	8
4		5	7	7
5			5	6
6			2	4
7				3
8				2

* Payoff for successful entrants as a function of "c"

The session details are described in Table 3.2.[34] In this experiment, there are eight sessions. In each session there are two blocks of twelve rounds. In one of the two blocks the rank is determined randomly, in the other block the rank depends on the skill, which is consistent with CL. This design feature is implemented in a within-subject design, i.e. the same subjects participate in both blocks of rounds. Subjects are told in advance in which block of rounds the rank is assigned randomly and in which it depends on their skill. In half of the sessions the block of rounds with random rank is run first, followed by the block of rounds with skill dependent rank. In the other half of the sessions the order is reversed, i.e. the block of rounds with skill dependent rank is run first to control for the order effect.

The recruitment procedure differed in information provided to the subjects before participating in the experiment. In the sessions one to four the subjects were recruited the usual way by saying that they were invited to participate in the experiment with an opportunity to make money. In sessions five to eight the subjects were also told before their participation (when recruited) that the payoff of the experiment would depend on

[34] These are identical to CL's experiment.

skill, especially knowledge about current events and sports.[35] In these sessions it was possible for subjects, who were confident of their abilities, to self-select themselves into the experiment.

Table 3.2 Description Of The Sessions

Session #	n	Recruitment Procedure	Block Order	Skill
1	12	No self-selection	R/S	Maze
2	14	No self-selection	S/R	Maze
3	16	No self-selection	R/S	Quiz
4	16	No self-selection	S/R	Quiz
5	16	Self-selection	R/S	Quiz
6	16	Self-selection	S/R	Quiz
7	14	Self-selection	R/S	Quiz
8	14	Self-selection	S/R	Quiz

R= random-rank, S=skill-rank

Table 3.3 shows the capacities used in each round. These are in the same sequence as in CL's experiment in order to maintain the procedures as close to the original experiment as possible. Along with their entry decision, subjects forecasted how many entrants they expect in that round. For each correct forecast they earned $1. CL used these forecasts to distinguish between subjects who entered because they underestimated the number of competitors from the subjects who entered because they thought they were above average, i.e. overconfident about their skills.

[35] The information included in the invitation email was as follows: "Earn money in an experiment in which performance on sports or current events trivia will determine your payoff. If you are very good you might earn a considerable sum of money."

Table 3.3 Market Capacity "c" Values

Round	Session 1	Session 2	Session 3-6	Session 7 and 8
1	2	8	2	4
2	4	4	6	2
3	8	2	4	6
4	6	6	4	8
5	4	4	2	6
6	2	2	6	4
7	8	8	4	2
8	6	6	6	8
9	4	4	2	6
10	6	2	6	4
11	8	8	4	2
12	2	6	2	8

I want to ask the same questions as CL but using data from my experiment. I thus test the following hypotheses:

Hypothesis 1: There is lower industry profit in skill-rank rounds than in random-rank rounds.

If subjects are overconfident, they will enter the market more often in skill-rank rounds, which will result in lower industry profits, i.e. if the number of entrants is higher than the c, the industry profit will be negative.

Hypothesis 2: The profit differentials between skill-rank and random-rank rounds in four sessions with self-selection are larger than in the remaining four sessions with no self-selection.

The larger the skill-random profit differential the more entry in the skill-rank rounds. If the entrants neglect the reference group, i.e. enter more because of the overconfidence in their skill but ignore that all the other entrants are doing the same thing, the differential between the skill-rank and random-rank will be larger in sessions with self-selection than in session with no self-selection.

Hypothesis 3: Expected average profit is smaller in skill-rank rounds than in random-rank rounds.

Expected average profit is calculated based on the forecasts of the subjects. If subjects decide to enter because they think fewer people will enter, then the expected average profit will be higher in skill-rank rounds than in random-rank rounds. If, however, subjects enter more often because they are more overconfident about their relative skills, the expected average profit will be lower in skill-rank rounds than in random-rank rounds (based on Equation 1 in Section 3.5.3 below).

A total of 118 subjects participated in the experiment. The experimental sessions were conducted in the New Zealand Experimental Economics Laboratory (NZEEL) at the University of Canterbury. Subjects were recruited using the online database system ORSEE(Greiner, 2004). Each subject only participated in a single session of the study, and had not participated in any similar market entry experiment at NZEEL. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). The number of subjects in a session varied from 12 to 16. All sessions were run under a single-blind social distance protocol.[36] On average, a session lasted 50 minutes including the payment and subjects earned 13.80 NZD.[37]

Upon entering the laboratory, subjects were asked to sit in a cubicle of their choice. At the beginning of the experiment instructions (provided in Appendix A) were handed out, as well as projected onto a screen and read aloud. Afterward subjects had a few minutes to go through the instructions again, this time in their own pace. Any questions arising were answered in private. All subjects had to answer control questions (provided in Appendix B) correctly, assessing their understanding of the instructions, before they could proceed to the decision making part of the experiment. After the control questions, subjects entered their decisions on market entry in each of 24 rounds and only after these decisions the subjects were asked to participate in a task, which determined their rank for the skill based rounds. Upon the completion of the experiment, they were also asked to fill out a questionnaire. Subjects were then called one by one to receive their payment in private in a payment room.

[36] In a single-blind treatment, decisions and payoffs are anonymous with respect to other subjects, but not to the experimenters.

[37] For reference, at the time of the experiment this was approximately 10.30 USD and the adult minimum wage in New Zealand was 14.25 NZD per hour.

3.5 Results

3.5.1 Skill-Rank Versus Random-Rank Rounds

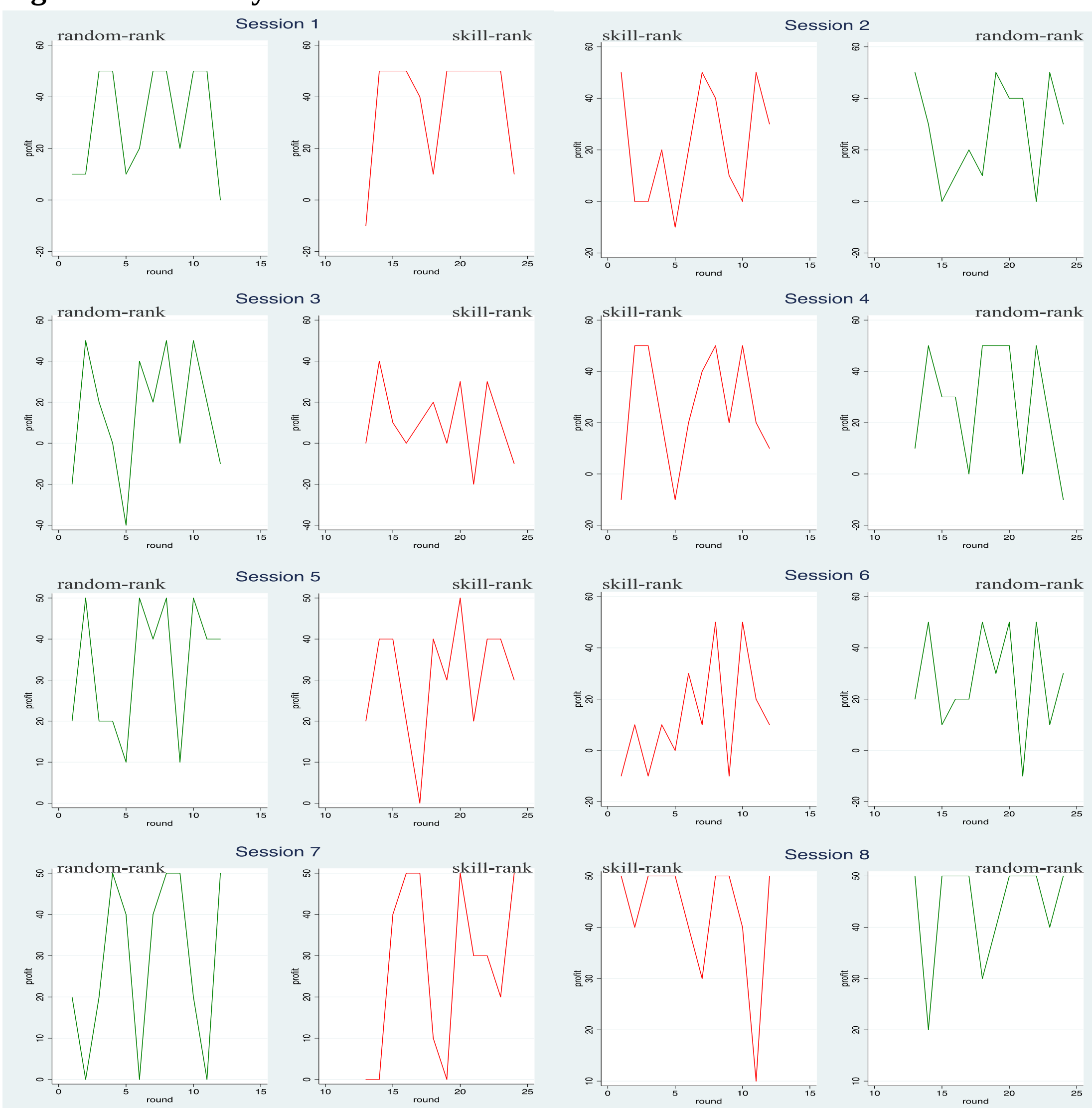
A game theoretic concept with a set of strategies, one for each player, such that each player's strategy is best for her, given that all other players are playing their equilibrium strategies is called Nash equilibrium.[38] Nash Equilibrium exists when there is no unilateral profitable deviation from any of the players involved. In other words, no player in the game would take a different action as long as every other player sticks to her strategy. The pure equilibrium of the market entry game is (c+5). With c=8, for example, a Nash equilibrium occurs when 13 players enter, with c=6, when 11 enter, with c=4, when 9 and with c=2 when 7 players enter. For example, if c=6 and 11 subjects enter, then an entrant has 1/11 probability of receiving $14, 1/11 probability of receiving $12, 1/11 probability of receiving $10, 1/11 probability of receiving $7, 1/11 probability of receiving $5 and 1/11 probability of receiving $2, because the top c entrants share $50 proportionally, and a 5/11 probability of losing $10. The expected payoff with 11 entrants thus is (1/11)*14+(1/11)*12+(1/11)*10+(1/11)*7+(1/11)*5+(1/11)*2+(5/11)*(-10)=$0.002. If one more player entered, the expected payoff would be (1/12)*14+(1/12)*12+(1/12)*10+(1/12)*7+(1/12)*5+(1/12)*2+(6/12)*(-10)=$-0.833. The player who entered could have avoided expected loss by staying out. If fewer than 11 players entered a player who stayed out could have received a positive payoff by entering, the expected payoff with 10 entrants is $1. For the mixed strategies (games in which players randomize), the expected payoff must be at least as large as that obtainable by any other strategy. There is a unique symmetric mixed-strategy equilibrium of the

[38] In 1950, John Nash's 28-page dissertation contained the definition and properties of Nash equilibrium.

market entry game in which (risk-neutral) players enter with a probability close to (c+5)/no of entrants (Camerer and Lovallo, 1999).

CL control for "risk attitude" by having subjects participate in both random- and skill-rank rounds, which gives an empirical estimate of observed equilibrium without having to impose any a priori assumptions about risk preferences. That is, since subjects participate in both random- and skill-rank rounds, their decisions in the random-rank rounds act as a within subject control for risk preference.

Figure 3.1 Industry Profits in Random-Rank and Skill-Rank Rounds



Moving onto the results from my experiment, Table 3.4 lists the total amount of money earned by subjects (depicted in Figure 3.1), i.e. the industry profit per round in each experimental session, by rank condition. Industry profit is calculated by adding the payoffs of successful entrants and payoffs (losses) of unsuccessful entrants in each round. In the majority of random-rank rounds (81/96 or 84%) industry profit is strictly positive. It is negative in 5 rounds (5% of the random-rank rounds). Average industry profit across rounds is $29.28. In the skill-rank rounds industry profit is strictly positive in 76/96 or 79% of rounds and negative in 9 rounds. Average profit across the skill-rank round is $26.14.

Table 3.4 Industry Profits by Rounds

Profit for random-rank rounds

		Rounds												
Session #	**n**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total**
1	12	10	10	50	50	10	20	50	50	20	50	50	0	370
2	14	50	30	0	10	20	10	50	40	40	0	50	30	330
3	16	-20	50	20	0	-40	40	20	50	0	50	20	-10	180
4	16	10	50	30	30	0	50	50	50	0	50	20	-10	330
5	16	20	50	20	20	10	50	40	50	10	50	40	40	400
6	16	20	50	10	20	20	50	30	50	-10	50	10	30	330
7	14	20	0	20	50	40	0	40	50	50	20	0	50	340
8	14	50	20	50	50	50	30	40	50	50	50	40	50	530

Profit for skill-rank rounds

		Rounds												
Session #	**n**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Total**
1	12	-10	50	50	50	40	10	50	50	50	50	50	10	450
2	14	50	0	0	20	-10	20	50	40	10	0	50	30	260
3	16	0	40	10	0	10	20	0	30	-20	30	10	-10	120
4	16	-10	50	50	20	-10	20	40	50	20	50	20	10	310
5	16	20	40	40	20	0	40	30	50	20	40	40	30	370
6	16	-10	10	-10	10	0	30	10	50	-10	50	20	10	160
7	14	0	0	40	50	50	10	0	50	30	30	20	50	330
8	14	50	40	50	50	50	40	30	50	50	40	10	50	510

Hypothesis 1 predicts lower industry profit (resulting from more entry) in skill-rank rounds than in random-rank rounds. The hypothesis is based on a conjecture that when subjects are betting on their own skill they will enter more often, which will in turn lower the total industry profit. Result 1 summarizes the finding.

Result 1: The industry profit is lower in skill-rank compared to random-rank rounds.

Support for Result 1: Following the original paper by CL, I also test for industry profit differences between skill-rank and random-rank rounds using a matched pair t-test in the following way. Recall that each experimental session consists of twelve random-rank and twelve skill-rank rounds. The order of ranks alternates between the sessions, i.e. in session one subject ranks were determined randomly in rounds 1-12 and by skill in rounds 13-24, whereas in session two subject ranks were determined by skill in rounds 1-12 and randomly in rounds 13-24. Industry profits by round in all sessions are depicted in Figure 3.1. I took the industry profit from the first twelve random-rank rounds in session one and matched them with industry profits from the first twelve skill-rank rounds in session two. Similarly, I took the industry profits from skill-rank rounds in session three and matched them with the industry profits from random-rank rounds in session four. In the same way I matched session five with six and seven with eight. Each pair of rounds being compared has the same value of c, takes place at the same experimental time and has the same history or path of previous values of c and differs only in how the rank was determined. Sessions one and two are exceptions, because of a different order of c. That is, subjects in session one face a different order of c than subjects in session two and thus these sessions are not directly comparable. I did this in order to replicate the original design by CL and so preserved the order of c's in sessions one and two. However, I exclude these two sessions from the matched pairs t-test (Table 3.4 reports industry profits for all sessions for completeness).[39] A matched pair two tailed t-test with these

[39] I do this because it is not appropriate to pair sessions one and two, as subjects face different capacities in the rounds to be paired.

comparisons suggests that there is more entry in the skill-rank rounds (i.e. overconfidence about one's skill) as the industry profits are lower in skill-rank than in random-rank rounds (p-value=0.084).

I also ran a matched pairs t-test with the total profits per session, as these are independent variables, which does not necessarily hold for the profits in individual rounds. In this case, there is no difference between the profits in random-rank and skill-rank rounds (p-value=0.527).

Attempting to replicate the study by CL I have used the same tests with preserving the same observation units. I recognize, however, that CL used their test assuming the independence of observations. A decision of a subject in round one, however, is not independent of a decision of a subject in round two since it is still the same subject making their decision about entering the market for 24 rounds. In order to eliminate this problem I ran a separate t-test for each round in sessions three to eight (again, excluding sessions one and two because of the different values of capacities). These observations are truly independent. The setback of this approach is a very small number of observations (n=3) and less statistical power but it does not violate the independency assumption. For this, I used the non-parametric Mann-Whitney test the results of which are presented in Table 3.5 for each round.

Table 3.5 Wilcoxon Signed-Rank Test for Differences in Industry Profits by Round

Round	**Wilcoxon signed-rank test (p-value)**
1, 5	0.786
2, 4, 6, 9, 11, 17	0.166
3, 8, 20, 22	0.103
7, 10, 13-16, 18, 19, 23	0.109
12, 24	0.285
21	0.782

In order to be able to use the data from sessions one and two (which do not have the same order and history path of c) I have calculated a normalised entry rate. The normalised entry rate is the ratio of the number of entrants and the actual capacity c in the respective round (presented in Table 3.6), where 100% means that the number of entering subjects was exactly the same as the c for a specific round. If the normalised entry rate is higher than 100%, it means that there were more people entering than c and if this is less than 100%, the market was not saturated and it was possible for more participants to enter the market and be profitable.

Table 3.6 The Normalised Entry Rates for Random-Rank and Skill-Rank Rounds

Normalised entry rate (%) for random-rank rounds

		Rounds												
Session #	**n**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Avg.**
1	12	300	200	13	33	200	250	38	50	175	33	38	350	140.0
2	14	100	150	350	167	175	300	75	117	125	350	75	133	176.4
3	16	450	83	175	225	550	117	175	100	350	100	175	400	241.7
4	16	300	67	150	150	350	50	100	100	350	100	175	400	191.0
5	16	250	67	175	175	300	83	125	83	300	67	125	150	158.3
6	16	250	100	200	175	250	100	150	83	400	100	200	200	184.0
7	14	175	350	150	88	117	225	150	75	83	175	350	50	165.7
8	14	100	250	33	38	50	150	150	13	67	100	150	38	94.9

Normalised entry rate (%) for skill-rank rounds

		Rounds												
Session #	**n**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Avg.**
1	12	400	75	38	33	125	300	38	67	100	67	38	300	131.8
2	14	63	225	350	150	250	250	88	117	200	350	100	133	189.7
3	16	350	117	200	225	300	150	225	133	450	133	200	400	240.3
4	16	400	100	100	175	400	150	125	83	250	83	175	300	195.1
5	16	250	117	125	175	350	117	150	100	250	117	125	200	173.0
6	16	400	167	250	200	350	133	200	100	400	100	175	300	231.3
7	14	225	350	117	63	100	200	350	75	133	150	250	50	171.9
8	14	75	150	67	63	83	125	200	13	67	125	300	38	108.8

By calculating the normalised entry rate I am able to control for different c's. The hypothesis tested is that there will be more entry (a higher normalised entry rate) in the skill-rank rounds than in the random-rank rounds. To test for differences in normalised entry rates between the skill-rank and random-rank rounds, I use a t-test, which does not detect a statistically significant difference (p-value=0.482)

I also tested for the difference in the normalised entry rate in skill-rank rounds and random-rank rounds both in self-selection sessions and no self-selection sessions. The Mann-Whitney test does not detect statistically significant differences in the normalised entry rate between skill- and random-rank rounds in sessions with self-selection (p-value=0.296) and with no self-selection (p-value=0.939). The result from these tests suggests that there is no higher entry in skill-rank rounds either in self-selection or in no self-selection sessions, i.e. subjects are not overconfident about their relative skills. Again, I ran an additional test with the average normalised entry rates, which control for the independence of observations. I decided to use a non-parametric Mann-Whitney test because of the small number of observations. The Mann-Whitney test does not detect a statistically significant difference in the average normalised entry rate between skill-rank and random-rank rounds in no self-selection sessions (8 observations, p-value=1.000) and or in the average normalised entry rates between skill-rank and random-rank rounds in sessions with self-selection (8 observations, p-value=0.387). This is in line with the previous tests on normalised entry rates with no difference between skill and random-rank rounds in self-selection sessions and in no self-selection sessions.

Similarly as with the industry profit, because of the independence of the observations, I ran the Mann-Whitney test (see Table 3.7) for each round in all eight sessions to test for differences in the normalised entry rate.[40] The drawback of this approach is small number of observations (n=4), but the advantage is that the independence assumption is not violated.

[40] There is no need to use Wilcoxon sign rank test as the c's are controlled for by using the normalized entry rate.

Table 3.7 The Mann-Whitney Test For Differences In The Normalised Entry Rates By Round

Round	The Mann-Whitney Test (p-value)
1,3	0.561
2, 21	0.773
4, 16, 17, 20, 23	0.885
5, 9	1.000
6	0.663
7	0.465
8, 24	0.559
10	0.309
11, 22	0.767
12	0.384
13	0.189
14	0.772
15	0.468
18	0.381
19	0.375

This complicated design allows me to test for within-subject comparisons as each subject participated both in random-rank and skill-rank rounds. Random-rank rounds have the exact same order of c's as skill-rank rounds, thus are directly comparable. To test if there is a difference in industry profit as well as in normalised entry rates between random-rank rounds and skill-rank rounds within a session I used the Mann-Whitney test; p-values for each session are reported in Table 3.8 below. Except for the industry profit in session 6, none of the results suggest that industry profits are different in random-rank rounds than in skill-rank round.

Table 3.8 The Mann-Whitney Test For The Difference In Industry Profits And Normalised Entry Rate Between Random-Rank And Skill-Rank Rounds

Session #	Industry profit (p-value)	Normalised entry rate (p-value)
1	0.476	0.771
2	0.500	0.728
3	0.560	0.602
4	0.788	0.794
5	0.592	0.622
6	0.077	0.307
7	0.953	0.954
8	0.713	0.643

3.5.2 Self-Selection Versus No Self-Selection Treatments

The average per-round industry profits reported in the first and second row of Table 3.9 are $25.21 and $23.75 for random- and skill-rank rounds respectively the no self-selection sessions (1-4), (no statistically significant difference, t-test, p-value=0.757).[41] In sessions with self-selection (5-8) average profit is $33.33 and $28.54 for random- and skill-rank rounds, respectively (no statistically significant difference, t-test, p-value=0.218).

Table 3.9 The Average per Round Industry Profit

Session	Rank Order	Avg Industry profit
1-4	Random-rank	$25.21
	Skill-rank	$23.75
5-8	Random-rank	$33.33
	Skill-rank	$28.54

[41] When sessions one and two are excluded from the test, the result still remains insignificant, t(46)=0.478 and p-value=0.635.

Hypothesis 2 states that the profit difference between skill-rank and random-rank rounds in the treatments with self-selection are larger than in the treatments with no self-selection. This hypothesis is based on a conjecture that reference group neglect produces larger profit differentials between skill-rank rounds and random-rank rounds in the sessions with self-selected subjects than in the sessions with no self-selected subjects.

Result 2: The profit differentials between skill-rank and random-rank rounds are not larger in sessions with self-selection than sessions with no self-selection.

Support for Result 2: A t-test with equal variances comparing the skill-random profit differentials between sessions 3-4 and 5-8 does not find support for the hypothesis that differentials are larger in sessions with self-selection than with no self-selection (p-value=0.782). In addition, I also used skill-random total profit differences to run a t-test, which does not detect a statistically significant difference (p-value=0.768) in the profit differences between self-selection and no self-selection sessions. This result is different to the one obtained by CL. For comparison, see Table 3.10. In CL reference group neglect produces a larger skill-random rank entry differential in sessions with self-selected subjects. A t-test comparing the skill-random profit differentials for matched rounds between session 1-4 and 5-8 strongly rejects the hypothesis that differentials are the same in sessions with and without self-selection (p-value < 0.001).

Table 3.10 A Comparison of the Average per Round Industry Profit and the Test Results

	CL	Replication
Avg. profit random-rank	\$16.87	\$29.28
Avg. profit skill-rank	\$-1.56	\$26.14
Matched pairs t-test	t = -7.43 p < 0.001	t = 1.755 p = 0.084
Avg. profit no self-selection, random rank	\$19.79	\$25.21
Avg. profit no self-selection, skill-rank	\$10.83	\$23.75
Avg. profit self-selection, random rank	\$13.96	\$33.33
Avg. profit self-selection, skill-rank	\$-13.13	\$28.54
t-test	t = -4.08 p < 0.001	t = 0.278 p = 0.782

Similarly as in the previous section, I used normalised entry rate instead of profit to test Hypothesis 2. Again, a t-test reports that the normalised entry rate skill-random rank differentials between self-selection and no self-selection sessions are not statistically significant (p-value=0.397). A t-test does not detect a statistically significant difference in the normalised entry rate between self-selection and no self-selection sessions either in random-rank rounds (p-value=0.110) or in skill-rank rounds (p-value=0.420).

This suggests that entry rates in random-rank rounds as well as in the skill-rank rounds are roughly the same in self-selection and no self-selection sessions. For both random-rank and skill-rank rounds it appears that the entry rate does not depend on the procedure used to recruit subjects. In other words, those who self-selected themselves into the experiment do not have higher entry rates than those who were recruited without knowing their payoff might depend on their skill. Similarly as with the industry profits in the previous section, this approach violates the independence assumption. In order to address this problem I ran the Mann-Whitney test with the average values of normalised entry rate per round, which are independent of each other. When looking at the averages of random-rank rounds, the Mann-Whitney test does not detect a difference between self-selection and no self-selection sessions in the normalised entry rate (8 observations, p-value=0.248). When considering only skill-rank rounds, again, the Mann-Whitney test does not detect a difference in normalised entry rate between self-selection sessions and no self-selection sessions (8 observations, p-value=0.387).

3.5.3 Expected Profit Differences In Skill And Random Rounds

The previous tests illustrate the effect of overconfidence on entry and demonstrate that self-selection does not make the effect stronger. These tests, however, do not control for all possible explanations. Excessive entry in the skill-rank rounds may not necessarily be due to overconfidence about skills, but due to the subjects underestimating how many subjects will enter (CL call this the "blind spot" hypothesis). If the number of expected entrants is underestimated, this influences the subjects' payoffs because they enter even though they should not have. In order to test this I saliently asked subjects to forecast how many entrants they think there will be in each round.[42] I then used these numbers to compute the profit that a subject expects the average entrant to earn. This has been calculated in a following way:

[42] The question I asked in the experiment is: "How many people (including yourself) do you expect to enter the market in this round?" If a subject forecasted the number of entrants correctly, $1 was added to his payoff in a respective round.

$$E_j(\pi_{ijt}) = (50\text{-}10*(F_{ijt} - c_{it}))/ F_{ijt,} \tag{1}$$

where $E_j(\pi_{ijt})$ is the expected average profit, F_{ijt} is the forecast of subject j used to calculate the profit that subject j expects the average entrant to earn, and c_{it} is the capacity in the particular round.

This approach allows me to separate overconfidence from incorrect estimates of others' entry. I test the hypothesis that the expected average profit is smaller in skill-rank rounds than in random-rank rounds. If subjects decide to enter in the skill-rank rounds because they think that fewer other subjects will enter, the expected average profit in the skill-rank rounds will be larger. Including $E_j(\pi_{ijt})$ in an entry regression will separate out the effect falsely attributed to skill. If, on the other hand, the subjects enter because they are more overconfident in the skill-rank rounds compared to random-rank rounds, not taking into account the number of entrants they expect to enter, the expected average profits will be smaller in skill-rank rounds than in random-rank rounds. The overconfident subjects will expect to earn more than the average entrant and enter even when the expected average profit is low.

Table 3.11 reports the difference between expected average profits in random-rank rounds (denoted π_r) and in skill-rank rounds (denoted π_s), using only the rounds in which subjects entered. Table 3.11 shows the mean difference π_r - π_s, averaged across entering subjects, the number and percentage of subjects who have a negative mean (i.e. who expect less average profit in skill-rank rounds), and the number and percentage of subjects whose expected average profit is negative, on average, across skill-rank rounds.

Table 3.11 The Average Differences In Expected Profits Per Entrant Between Random And Skill Rounds

Measure	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8	Total
Πr - Πs	-1.126	-0.665	0.023	1.832	-1.094	-0.886	-1.718	4.483	1.036
# of S's with Πr - Πs<0 (percent)	5/9 (55.6%)	6/14 (42.9%)	7/13 (53.8%)	4/12 (33.3%)	9/11 (81.8%)	6/14 (42.9%)	8/13 (61.5%)	2/8 (25%)	47/94 (50%)
# of S's with Πr <0 (percent)	0/9 (0%)	0/13 (0%)	3/12 (25%)	1/11 (9%)	0/10 (0%)	0/12 (0%)	0/11 (0%)	0/8 (0%)	4/86 (4.7%)

The mean difference π_r - π_s is negative in sessions one and two and positive in sessions three and four. A majority of subjects expect to earn more in skill-rank rounds. In session one 44% of the subjects expect to earn less in skill-rank rounds than in random-rank rounds. In session two, three and four, it is 50%, 46% and 67%, respectively.

In the self-selection sessions (5-8), the mean difference π_r - π_s is negative in all sessions, except for session eight. Subjects expect to earn more in skill-rank rounds than in random-rank rounds. In session five only 18% of the subjects expect to earn less in the skill-rank rounds. In session six, seven and eight it is 57%, 38% and 75%, respectively. A t-test does not detect a difference in the average difference of expected profits per person between the self-selection treatments and the no self-selection treatments at a 10% significance level (p-value=0.913).

Result 3: The expected average profit is not smaller in skill-rank rounds than in random-rank rounds.

Support for Result 3: A t-test comparing the expected average profit per entrant in skill-rank rounds and random-rank rounds does not provide evidence that these profits are different (p-value=0.879).

3.5.4 Logistic Regression

The size and significance of all variables' effects at once can be tested using a fixed effects logistic regression with entry decision as independent variable (Table 3.12) and all the variables that could be influencing the independent variable, i.e. c, the expected average profit $E_j(\pi_{ijt})$, skill-rank dummy (Skill=1), subject pool dummy (students studying economics or business economics, ECON=1), self-selection condition dummy (sessions with Self-selection=1), gender (Male=1), dummy for sessions with the maze (Maze=1), the order of rounds (Random/Skill=1 if random-rank rounds were run first) and interaction variable Self-selection*Skill. The interaction term was included in the regression to control for the situation when Self-selection and Skill interact. This means the effect of Skill on the entry decision might depend on whether the self-selected recruitment procedure was used. Theory suggests that subjects, who self-selected themselves into the experiment based on the information they received in the recruitment email, will enter more in the skill-rounds. Expected average profit, Skill-rank, ECON, Self-selection, Male and Maze are all statistically significant. The results suggest that the higher the expected profit, the fewer subjects enter. Also, ECON students and those who self-selected themselves into the experiment seem to enter less often. It appears that male subjects entered more often than female subjects and subjects in sessions with the maze entered more often than in the sessions with the quiz.

Table 3.12 Fixed Effects Logistic Regression Of Entry Decision, Sessions 1-8, n=2832, Panel Data[43]

Variable	Coefficient (Robust Std Err)	Marginal	z-statistic (p-value)
Intercept	-0.410 (0.237)		-1.73 (0.084)
c	-0.031 (0.023)	-0.007	-1.33 (0.182)
$E(\pi_{ijt})$	-0.014*** (0.004)	-0.003	-3.28 (0.001)
Skill	0.202* (0.109)	0.046	1.85 (0.064)
ECON	-0.447*** (0.084)	-0.103	-5.29 (0.001)
Self-selection	-0.246** (0.124)	-0.057	-1.99 (0.047)
Male	0.491*** (0.079)	0.113	6.22 (0.001)
Maze	0.325*** (0.118)	0.075	2.76 (0.006)
Order Random/Skill	0.128 (0.079)	0.030	1.63 (0.104)
Self-selection*Skill	-0.122 (0.157)		-0.78 (0.437)
Round dummy	yes		
Log-likelihood	-1847.55		Pseudo R2=0.035

Run on StataSE 12.0. Robust standard errors used. Round 24 is the omitted control variable.

*, **, *** refer to statistical significance at the 10%, 5% and 1% levels, respectively.

[43] Standard errors are not clustered to session level because of a small number of sessions.

3.5.5 Additional Analyses: Forecasts

Recall that subjects had to forecast number of entering subjects in each round. If the forecast was correct, an additional $1 was added to subject's payoff. Analysing the forecasting behaviour allows to distinguish between two types of over-entering the market. Over-entering can occur either because subjects forecast the amount of competition to be smaller than it actually is or because subjects are overconfident about their relative skill and decide to enter too often. There are many studies showing that forecasts are usually biased and violate rationality of expectations (Camerer, Kagel, and Roth, 1995). Forecasts are usually correlated with observable variables and usually follow an adaptive process in which forecast changes are related to past forecast errors (Nerlove, 1958).

Violations of rationality of naturally occurring forecasts could be due to Bayesian learning in an economy where the statistical process generating outcomes keeps changing (Caskey, 1985; Lewis, 1989). To control for changes in statistical process generating outcomes, several experiments examined forecasts of outcomes of a statistical process that is unknown to subjects but fixed throughout the experiment, and known to be fixed (Bolle, 1988; Garner, 1982; Schmalensee, 1976). Their results are generally inconsistent with rationality of expectations too, but suggest some learning and rationality in special setting. A notable exception is (Daniels and Plott, 1988) who studied forecasts in goods markets with price inflation that was induced by shifting supply and demand curves upward by 15% each round. Regressions indicated that subjects' forecasts were rational rather than adaptive.

Figure 3.2 Forecasted And Actual Number Of Entrants in Random-Rank Rounds, Averages

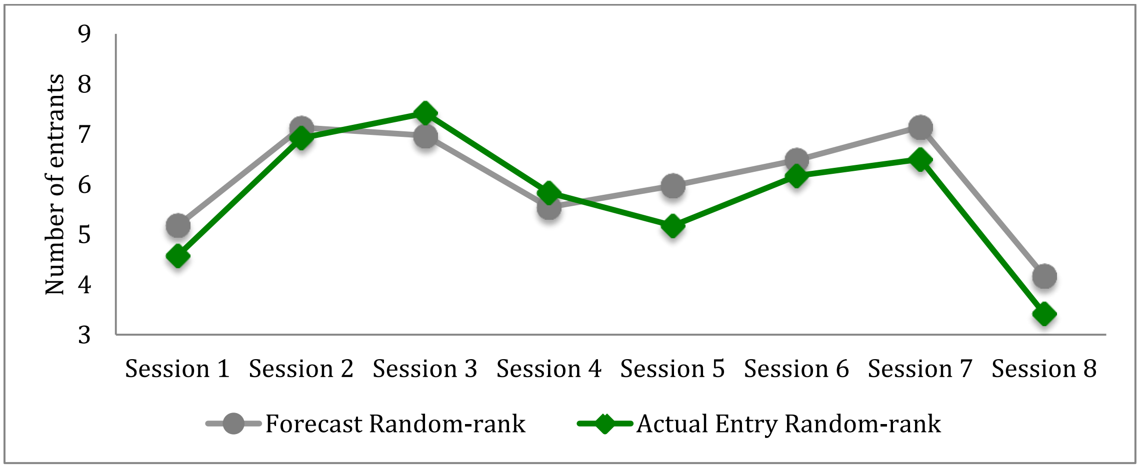


Figure 3.3 Forecasted And Actual Number Of Entrants in Skill-Rank Rounds, Averages

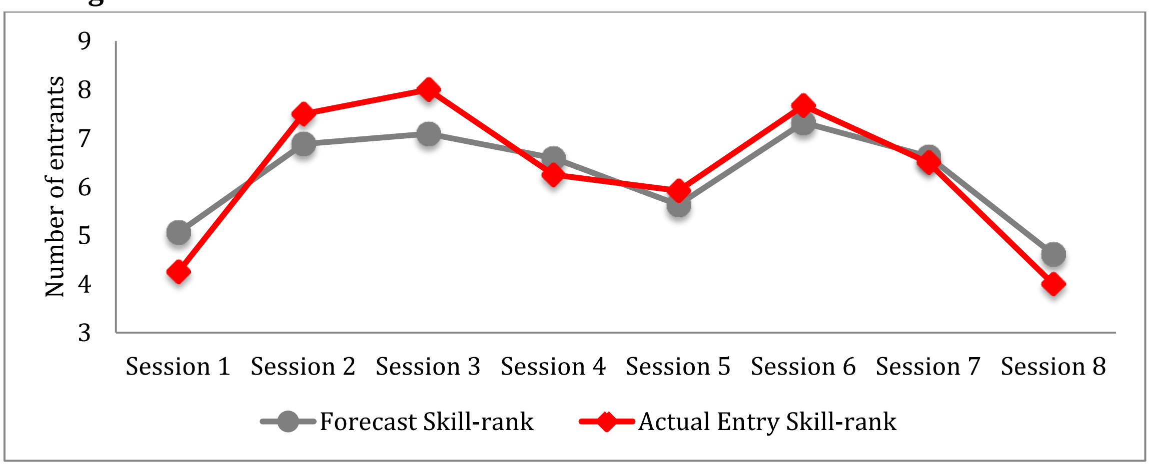
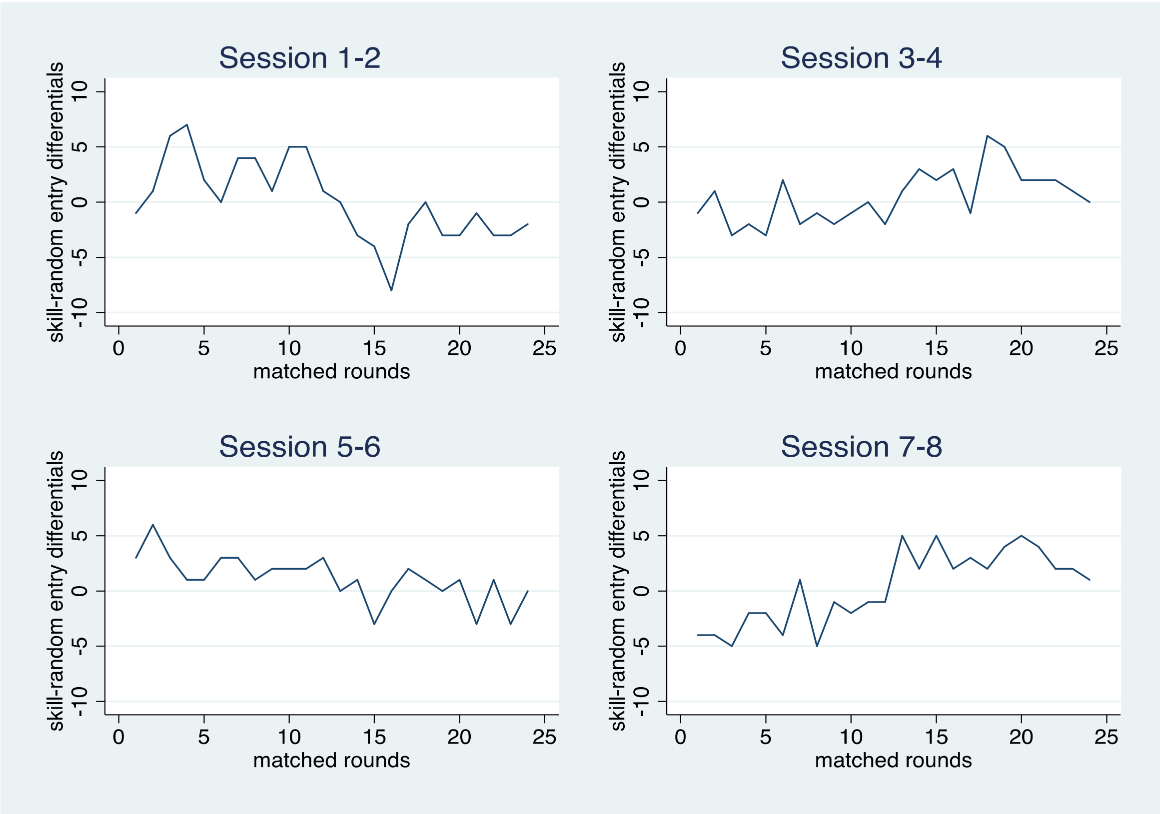


Figure 3.2 and Figure 3.3 present the averages of number of expected and actual entrants in random-rank rounds and skill-rank rounds in each session. On average, the number of forecasted entrants in all sessions in random-rank rounds is 6.07 subjects and 6.23 subjects in skill-rank rounds. The actual number of entrants in all sessions is on average 5.75 and 6.26 for random-rank rounds and skill-rank rounds, respectively. The difference between forecasted and actual entrants in random-rank rounds is not

statistically significantly different at the 10% significance level (Mann-Whitney p-value=0.599). In skill-rank rounds this difference is not statistically significant either (Mann-Whitney p-value=0.916). In random-rank rounds subjects forecast about 0.32 entrants too high and in skill-rank rounds they forecast is converging to the actual number of entrants.

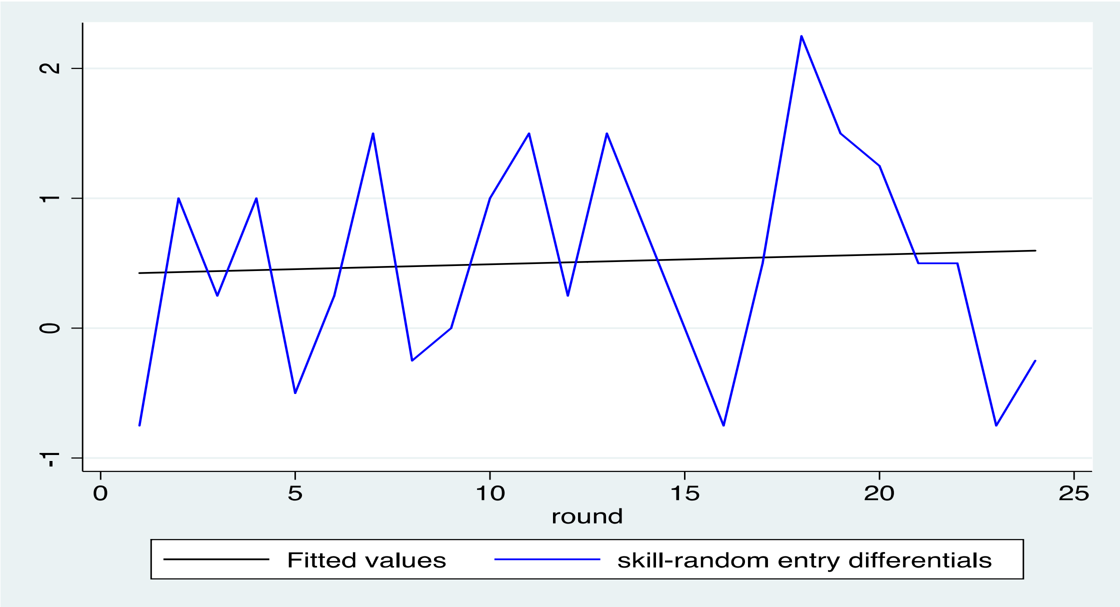
Figure 3.4 Matched-Pair Skill-Random Differentials In Number Of Entrants



The time series of matched-pair skill-random differentials in number of entrants has a slight downward trend across rounds in sessions 1-2 and 5-6 and slight upward trend in sessions 3-4 and 7-8 (Figure 3.4). My data thus shows mixed trends.

When the number of entrants is averaged across rounds, the total skill-random differentials in entry have a slight upward trend as opposed to CL's downward trend (Figure 3.5).

Figure 3.5 Averaged Skill-Random Entry Differentials



3.6 Demographics

As demographic data were collected in the post-experiment questionnaire I can test if any of these are influencing subjects' entry decisions. All variables that were potentially relevant were used in the regression. The results from an OLS regression are reported in Table 3.13 below.

NUMBER OF EXPECTED ENTRANTS is the number that each participant forecasted would enter the market in a respective round. SELF-SELECTION is a dummy variable that takes on the value of 1 for sessions 5-8, in which subjects were told in the recruitment process that their payoff will depend on their skill (current events and sports quiz). SKILL RANK is a dummy variable that takes on the value of 1 for rounds, in which the rank depends on the skill, i.e. 12 rounds in each session. MALE is a dummy variable that equals one when a subject reported their gender as male and 0 if female. NON NZ is a dummy variable that equals one when subjects selected a nationality of a country that was not New Zealand. Out of this group approximately 78.13% of Non New Zealanders chose an Asian nationality, and approximately 6.25% nominated a European nationality. ECON is a dummy variable that is 1 when subjects nominated an area of study that was 'Economics', or 'Business Economics'. LIVE WITH OTHERS is the number of people that currently live in the subject's household, MONEY is the amount of dollars that subjects nominated as their monthly non-accommodation budget, FINANCE STUDY is the proportion that subjects nominated as the fraction of their monthly budget that they fund themselves.

Table 3.13 Demographics Analysis, IN As Dependent Variable

	OLS Coefficient (p-value) Decision "IN"
NUMBER OF EXPECTED ENTRANTS	0.023*** (0.001)
SELF-SELECTION	-0.090*** (0.001)
SKILL RANK	0.031* (0.082)
MALE	0.111*** (0.001)
AGE	-0.015*** (0.001)
NON NZ	-0.033* (0.181)
SIBLINGS	-0.031**** (0.001)
ECON	-0.079*** (0.001)
RELATIVE INCOME	-0.020* (0.100)
CITY SIZE	0.103*** (0.001)
LIVE WITH OTHERS	0.026*** (0.001)
MONEY	0.001 (0.565)
FINANCE STUDY	0.001 (0.734)
NO. OF PEOPLE KNOWN	0.010 (0.362)
RELY	-0.002 (0.605)
CONSTANT	0.193** (0.020)

Run on StataSE 12.0. *, **, *** refer to statistical significance at the 10%, 5% and 1% levels, respectively.

RELATIVE INCOME is the income of subjects' parents in comparison to other families in New Zealand when subjects were 16 years of age. The larger the stated family income, the larger is the variable RELATIVE INCOME. When answering the question about RELATIVE INCOME, subjects had five possibilities to choose from, i.e. far below average, below average, average, above average, far above average. CITY SIZE is the size of the community where the subject has lived the most time of his life. Subjects could choose from the following four options: up to 2 000 inhabitants, 2 000 to 10 000 inhabitants, 10 000 to 100 000 inhabitants, more than 100 000 inhabitants. The more inhabitants they state the higher is the variable CITY SIZE. NO. OF PEOPLE KNOWN is the number subject knows in the session. RELY is a 9 point measure of how much a subject thinks the experimenter should rely on their data, where 9 is the most reliable.

MONEY and FINANCE STUDY were included as measures to control for income. NUMBER OF EXPECTED ENTRANTS has positive and significant effect on the decision to enter the market. This suggests that the more people the participants expect to enter, the more they tend to enter the market. Interestingly, the coefficient and the p-value of SELF-SELECTION indicates that in the self-selection sessions (5-8) the participants tend to enter the market less than in the sessions with no self-selection (sessions 1-4). In the SKILL RANK rounds the participants tend to enter the market more often than in the random rank rounds. NON NZ, MONEY, FINANCE STUDY, NO OF PEOPLE KNOWN and RELY are not significant at the 5% significance level. The variable ECON was used to control for the behaviour of economics students. Economics students are likely to have learnt about concepts such as rationality and backwards induction, and therefore tend to be more likely to play strategies closer to what standard game theory would predict. An interesting result of this analysis of demographics is the statistically significant coefficient related to the variable LIVE WITH OTHERS. This suggests that the more people in the same household the higher the frequency of entering the market.

3.7 Problems Arising In The Replication

Replicating an experiment can be challenging as one has to understand in detail someone else's ideas, what has been done, how it has been done and why. I came across a few issues I would like to address here. First of all, CL put a lot of emphasis on matching the industry profits from skill-rank rounds from one session with profits from random-rank rounds from another session in order to perform a matched pairs t-test. In this situation the order of capacities and the order of random vs. skill-ranks is important. CL compare session one with session two, three with four, five with six and seven with eight. All these session pairs have the same order of capacities except for the session one-session two pair. CL do not explain how they matched these two sessions or why it was possible to match them.

Later on, when seeking for an answer to the question if the profit differentials between skill-rank and random-rank rounds in four treatments with self-selection are larger than in the remaining four treatments with no self-selection, CL assume a normal distribution and use the t-test. The distribution, however, is not normal in skill-rank profits (Skewness and kurtosis test for normality, p-value=0.001) and random-rank profits (p-value=0.030). The data publicly available from this experiment was not complete, i.e. the data from two sessions were missing. I emailed prof. Camerer asking for missing data and clarification of some design aspects and tests used. I did not hear back from him so I phoned prof. Lovallo but I was not successful in reaching him after placing a call each day for 5 days. Then I sent an email to prof. Lovallo, to which he replied and told me where to find the instructions for the experiment. My remaining three emails about the design aspects and statistical tests remained unanswered. In Experimental Economics it is paramount to make exact and detailed description of procedures available in order to make the experiment available for later replications. As the exact procedures of CL experiment are not publicly available I was following them to the extent that was available from the paper.

3.8 Conclusion

One of the explanations for possible business failure is overconfidence of those who decide to enter the market. In this experiment I test whether overconfidence about one's skill influences the frequency of market entry decisions. I decided to replicate the experiment by CL, in which entrants' payoffs depend on their skills. I have obtained the same results as CL but only to some extent. I found support for the Hypothesis 1 that the industry profit is lower in skill-rank rounds than in random-rank rounds. The result from a matched pairs t-test suggests that there is more entry in the skill-rank rounds, resulting in lower industry profits in skill-rank than in random-rank rounds. This result suggests that subjects are overconfident about their skills. The profit differentials between skill-rank and random-rank rounds in the treatments with self-selection are not different from the differentials in the treatments with no self-selection, thus I did not find support for my Hypothesis 2. The above-mentioned tests control for the capacity by using a matched pairs t-test. Since the capacities does not match up in sessions one and two, I had to exclude them from my data analysis. In order to make use out of the data from sessions one and two I have calculated a normalised entry rate, which controls for capacity. I then used this normalised entry rate to run my tests. Using this data I did not find support for my Hypothesis 1 or 2, i.e. normalised entry rate was not higher in skill-rank than in random-rank rounds and skill-random differentials are not higher in sessions with self-selection than in sessions with no self-selection. As subjects forecasted expected number of entrants, I was able to test whether expected average profit is smaller in skill-rank than in random-rank rounds and found no statistically significant difference in expected average profits between these two ranks.

There might be more reasons to why I did not obtain the same results as CL. First of all, instead of using only male subjects in session three to eight, I used males and females in all of my sessions. Deaves, Luders, and Luo (2009) test among other things for gender effect in an experiment on overconfidence and they find little evidence that gender influences trading activity. Second, as detailed procedures were not available, it is possible that my procedures deviated from the original procedures to some extent. Finally, the subject pool was different as well (New Zealand students vs. US students). My results show that the overconfidence effect is very sensitive to experimental conditions.

Appendix A. Subject Instructions

Session 1 Maze Rounds First Random Rounds Second

INSTRUCTIONS

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Anonymity

The identity of the participants will not be revealed to other participants at any time during the experiment.

Show-up Fee

If you agree to participate in the experiment you will be given $5, which is yours to keep.

Structure of the Experiment

This experiment is computerized. If you have any problems entering your decision, please alert the experimenter. The experiment involves two sets of decisions. Each set of decisions consists of 12 rounds (i.e. 24 rounds in total). These two sets differ in how the rank is determined. In the first 12 rounds your rank will be determined by your speed of finishing the mazes (as will be explained later). In the second 12 rounds your rank will be determined randomly.

In each round you are asked to decide whether to enter the market or not. In the beginning of each round the market capacity “c” for that round will be announced. You can think of “c” as the size of the market. You will also be informed about the number of entrants in the previous round.

Decision Making Task

In each round you start with $10.

If you decide not to enter the market, you earn nothing and lose nothing; your earnings for that round will be $10.

If you decide to enter the market, your payoff in each round will depend on your rank relative to the ranks of other participants who entered the market and on the capacity “c”.

If you entered the market

Your rank and the capacity for that round determine if you are a successful or unsuccessful entrant. If your rank is less than or equal to the capacity, then you are a successful entrant. If your rank is greater than the capacity, then you are an unsuccessful entrant. The unsuccessful entrant will lose the $10 (s)he was given in the beginning of that round. The payoffs of successful entrants as a function of “c” are shown in the table below.

ONE of all 24 rounds will be chosen randomly and your rank and decision in this round will determine your payoff.

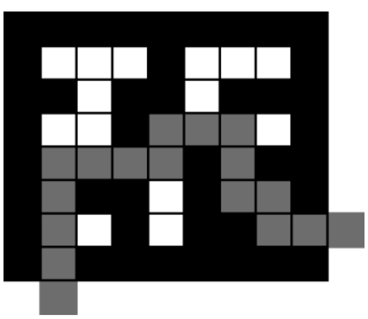
Rank	Capacity “c”			
	2	4	6	8
1	33	20	14	11
2	17	15	12	10
3		10	10	8
4		5	7	7
5			5	6
6			2	4
7				3
8				2

All participants will take part in both sets of decisions in the same order. In each round you will be also asked to estimate the number of people (including you) that you expect to enter the market in that round. If your estimation of the number of entrants is the same as the actual number of entrants in that round, additional $1 will be added to your payoff in that round.

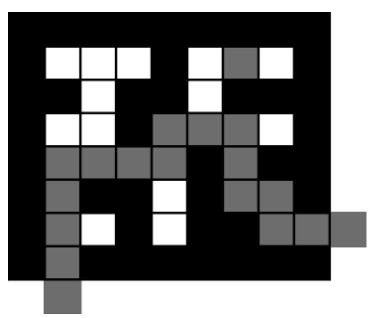
The Maze

After you finish all 24 rounds in the decision making task, you will be given five mazes to solve. You need to find the shortest way from one end of the maze to another. If you have highlighted all the correct squares in the maze, the OK button will pop up. Click OK in order to continue. The participant, who finishes the mazes the fastest, will be ranked number 1. A participant, who is the second fastest, will be ranked number 2 and so on.

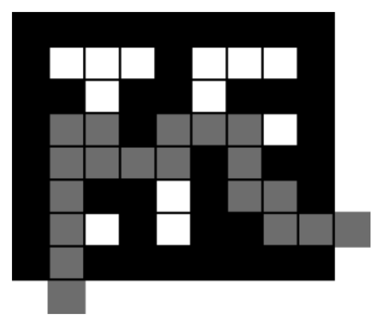
Correct way:



Incorrect way:



Incorrect way:



Example

Suppose “c” is 2 and four participants decide to enter the market. The entrant with rank number 1 earns $33 and the entrant with rank number 2 earns $17. The entrants with rank number 3 and number 4 lose $10, i.e. their payoff for that round will be 0.

Payment of Experiment Earnings

ONE of all 24 rounds will be chosen randomly and your rank and decisions in this round will determine your payoff.

All money will be paid to you in cash at the end of the experiment. Because your decision is private, we ask that you do not tell anyone your decision or your earnings either during or after the experiment. We also ask you to not gather near the lab after you receive your payment.

Are there any questions?

Session 2 Random Rounds First Maze Rounds Second

INSTRUCTIONS

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Anonymity

The identity of the participants will not be revealed to other participants at any time during the experiment.

Show-up Fee

If you agree to participate in the experiment you will be given $5, which is yours to keep.

Structure of the Experiment

This experiment is computerized. If you have any problems entering your decision, please alert the experimenter. The experiment involves two sets of decisions. Each set of decisions consists of 12 rounds (i.e. 24 rounds in total). These two sets differ in how the rank is determined. In the first 12 rounds your rank will be determined randomly. In the second 12 rounds your rank will be determined by your speed of finishing the mazes (as will be explained later).

In each round you are asked to decide whether to enter the market or not. In the beginning of each round the market capacity "c" for that round will be announced. You can think of "c" as the size of the market. You will also be informed about the number of entrants in the previous round.

Decision Making Task

In each round you start with $10.

If you decide not to enter the market, you earn nothing and lose nothing; your earnings for that round will be $10.

If you decide to enter the market, your payoff in each round will depend on your rank relative to the ranks of other participants who entered the market and on the capacity “c”.

If you entered the market

Your rank and the capacity for that round determine if you are a successful or unsuccessful entrant. If your rank is less than or equal to the capacity, then you are a successful entrant. If your rank is greater than the capacity, then you are an unsuccessful entrant. The unsuccessful entrant will lose the $10 (s)he was given in the beginning of that round. The payoffs of successful entrants as a function of “c” are shown in the table below.

ONE of all 24 rounds will be chosen randomly and your rank and decision in this round will determine your payoff.

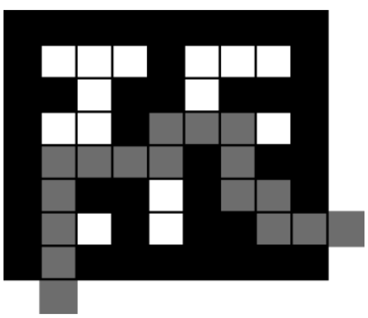
Rank	Capacity “c”			
	2	4	6	8
1	33	20	14	11
2	17	15	12	10
3		10	10	8
4		5	7	7
5			5	6
6			2	4
7				3
8				2

All participants will take part in both sets of decisions in the same order. In each round you will be also asked to estimate the number of people (including you) that you expect to enter the market in that round. If your estimation of the number of entrants is the same as the actual number of entrants in that round, additional $1 will be added to your payoff in that round.

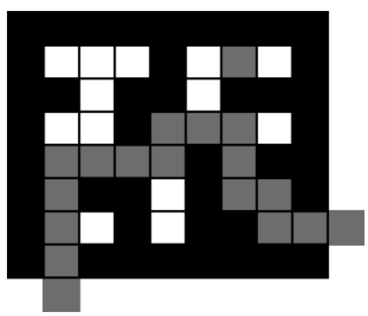
The Maze

After you finish all 24 rounds in the decision making task, you will be given five mazes to solve. You need to find the shortest way from one end of the maze to another. If you have highlighted all the correct squares in the maze, the OK button will pop up. Click OK in order to continue. The participant, who finishes the mazes the fastest, will be ranked number 1. A participant, who is the second fastest, will be ranked number 2 and so on.

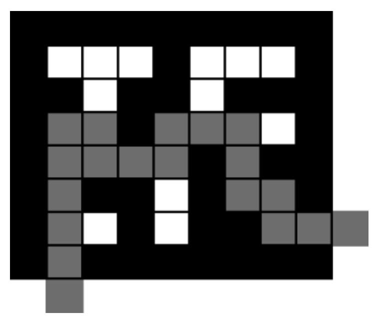
Correct way:



Incorrect way:



Incorrect way:



Example

Suppose “c” is 2 and four participants decide to enter the market. The entrant with rank number 1 earns $33 and the entrant with rank number 2 earns $17. The entrants with rank number 3 and number 4 lose $10, i.e. their payoff for that round will be 0.

Payment of Experiment Earnings

ONE of all 24 rounds will be chosen randomly and your rank and decisions in this round will determine your payoff.

All money will be paid to you in cash at the end of the experiment. Because your decision is private, we ask that you do not tell anyone your decision or your earnings either during or after the experiment. We also ask you to not gather near the lab after you receive your payment.

Are there any questions?

Sessions 3, 5, 7 Random Rounds First Quiz Rounds Second

INSTRUCTIONS

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Anonymity

The identity of the participants will not be revealed to other participants at any time during the experiment.

Show-up Fee

If you agree to participate in the experiment you will be given $5, which is yours to keep.

Structure of the Experiment

This experiment is computerized. If you have any problems entering your decision, please alert the experimenter. The experiment involves two sets of decisions. Each set of decisions consists of 12 rounds (i.e. 24 rounds in total). These two sets differ in how the rank is determined. In the first 12 rounds your rank will be determined randomly. In the second 12 rounds your rank will be determined by your score on a quiz (as will be explained later).

In each round you are asked to decide whether to enter the market or not. In the beginning of each round the market capacity "c" for that round will be announced. You can think of "c" as the size of the market. You will also be informed about the number of entrants in the previous round.

Decision Making Task

In each round you start with $10.

If you decide not to enter the market, you earn nothing and lose nothing; your earnings for that round will be $10.

If you decide to enter the market, your payoff in that round will depend on your rank relative to the ranks of other participants who entered the market and on the capacity "c".

If you entered the market

Your rank and the capacity for that round determine if you are a successful or unsuccessful entrant. If your rank is less than or equal to the capacity, then you are a successful entrant. If your rank is greater than the capacity, then you are an unsuccessful entrant. The unsuccessful entrant will lose the $10 (s)he was given in the beginning of that round. The payoffs of successful entrants as a function of "c" are shown in the table below.

ONE of all 24 rounds will be chosen randomly and your rank and decision in this round will determine your payoff.

Rank	Capacity "c"			
	2	4	6	8
1	33	20	14	11
2	17	15	12	10
3		10	10	8
4		5	7	7
5			5	6
6			2	4
7				3
8				2

All participants will take part in both sets of decisions in the same order. In each round you will be also asked to estimate the number of people (including you) that you expect to enter the market in that round. If your estimation of the number of entrants is the same as the actual number of entrants in that round, additional $1 will be added to your payoff in that round.

The Quiz

After you finish all 24 rounds in the decision making task, you will be asked to participate in a multiple choice quiz. There are 30 sports & current events questions in the quiz, each question has only one correct answer. You will have 10 minutes to answer all questions. A participant with the most correct answers will be ranked number 1, etc. If two or more participants correctly answered the same number of questions, the ties will be broken by the shorter amount of time taken to answer all questions.

Example

Suppose "c" is 2 and four participants decide to enter the market. The entrant with rank number 1 earns $33 and the entrant with rank number 2 earns $17. The entrants with rank number 3 and number 4 lose $10, i.e. their payoff for that round will be 0.

Payment of Experiment Earnings

ONE of all 24 rounds will be chosen randomly and your rank and decisions in this round will determine your payoff.

All money will be paid to you in cash at the end of the experiment. Because your decision is private, we ask that you do not tell anyone your decision or your earnings either during or after the experiment. We also ask you to not gather near the lab after you receive your payment.

Are there any questions?

Sessions 4, 6, 8 Quiz Rounds First Random Rounds Second

INSTRUCTIONS

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question after we finish reading the instructions, please raise your hand and the experimenter will approach you and answer your question in private.

Anonymity

The identity of the participants will not be revealed to other participants at any time during the experiment.

Show-up Fee

If you agree to participate in the experiment you will be given $5, which is yours to keep.

Structure of the Experiment

This experiment is computerized. If you have any problems entering your decision, please alert the experimenter. The experiment involves two sets of decisions. Each set of decisions consists of 12 rounds (i.e. 24 rounds in total). These two sets differ in how the rank is determined. In the first 12 rounds your rank will be determined by your score on a quiz (as will be explained later). In the second 12 rounds your rank will be determined randomly.

In each round you are asked to decide whether to enter the market or not. In the beginning of each round the market capacity "c" for that round will be announced. You can think of "c" as the size of the market. You will also be informed about the number of entrants in the previous round.

Decision Making Task

In each round you start with $10.

If you decide not to enter the market, you earn nothing and lose nothing; your earnings for that round will be $10.

If you decide to enter the market, your payoff in that round will depend on your rank relative to the ranks of other participants who entered the market and on the capacity "c".

If you entered the market

Your rank and the capacity for that round determine if you are a successful or unsuccessful entrant. If your rank is less than or equal to the capacity, then you are a successful entrant. If your rank is greater than the capacity, then you are an unsuccessful entrant. The unsuccessful entrant will lose the $10 (s)he was given in the beginning of that round. The payoffs of successful entrants as a function of "c" are shown in the table below.

ONE of all 24 rounds will be chosen randomly and your rank and decision in this round will determine your payoff.

Rank	Capacity "c"			
	2	4	6	8
1	33	20	14	11
2	17	15	12	10
3		10	10	8
4		5	7	7
5			5	6
6			2	4
7				3
8				2

All participants will take part in both sets of decisions in the same order. In each round you will be also asked to estimate the number of people (including you) that you expect to enter the market in that round. If your estimation of the number of entrants is the same as the actual number of entrants in that round, additional $1 will be added to your payoff in that round.

The Quiz

After you finish all 24 rounds in the decision making task, you will be asked to participate in a multiple choice quiz. There are 30 sports & current events questions in the quiz, each question has only one correct answer. You will have 10 minutes to answer all questions. A participant with the most correct answers will be ranked number 1, etc. If two or more participants correctly answered the same number of questions, the ties will be broken by the shorter amount of time taken to answer all questions.

Example

Suppose "c" is 2 and four participants decide to enter the market. The entrant with rank number 1 earns $33 and the entrant with rank number 2 earns $17. The entrants with rank number 3 and number 4 lose $10, i.e. their payoff for that round will be 0.

Payment of Experiment Earnings

ONE of all 24 rounds will be chosen randomly and your rank and decisions in this round will determine your payoff.

All money will be paid to you in cash at the end of the experiment. Because your decision is private, we ask that you do not tell anyone your decision or your earnings either during or after the experiment. We also ask you to not gather near the lab after you receive your payment.

Are there any questions?

Appendix B. Control Questions

1. How much would you earn in a round if c=6, you entered and your rank was 5 among the entrants?

2. How much would you earn in a round if c=2, you entered and your rank was 4 among the entrants?

3. How much would you earn in a round if you decided not to enter the market?

4. How many rounds are there in total in this experiment?

Appendix C. The Quiz

1. Which of the following countries is no longer in the hunt for a spot at the 2015 Rugby World Cup?

    A. Zimbabwe
    B. Uruguay
    C. Hong Kong
    D. **Madagascar**

2. How many consultants are being hired to help the NFL deal with domestic violence issues?

    A. 1
    B. 2
    C. **3**
    D. 4

3. Who won the Wimbledon men's title in 2008?

    A. Roger Federer
    B. **Rafael Nadal**
    C. Novak Djokovic
    D. Andy Murray

4. Which famous yacht race starts on Boxing Day each year?

    A. Transpacific Yacht Race
    B. Newport to Bermuda
    C. **The Sydney to Hobart Yacht Race**
    D. Regata del Sol al Sol

5. How many gold medals did Michael Phelps win at the 2008 Olympics?

    A. **Eight**
    B. Seven
    C. Ten
    D. Nine

6. How many red balls are there in the snooker?

    A. **15**
    B. 16
    C. 17
    D. 18

7. Who became the first non-European to win the Tour de France cycle race?

    A. Cadel Evans
    B. **Greg Lemond**
    C. Lance Armstrong
    D. Geraint Thomas

8. When did the first Olympic mascot appear?

    A. 1956
    B. **1972**
    C. 1980
    D. 1984

9. Who has been named honorary president of World Cup organisers England Rugby 2015?

   A. Prince Charles
   B. Prince William
   C. **Prince Harry**
   D. Prince Andrew

10. Who has been appointed as the United Nations representative on climate change?

   A. Brad Pitt
   B. George Clooney
   C. **Leonardo DiCaprio**
   D. Russell Crowe

11. Which city did 'Vogue' recently name as the Coolest Little City in the World?

   A. **Wellington**
   B. Nelson
   C. Christchurch
   D. Taupo

12. How long was the longest successful basketball shot?

   A. **33.45 m**
   B. 43.45 m
   C. 47.45 m
   D. 49.45 m

13. In what year was Matthew Ridge first named in the All Blacks?

A. **1989**
B. 1991
C. 1993
D. 1995

14. Which book won the top New Zealand Post Book Award this year?

A. Eleanor Catton: The Luminaries
B. **Peter McLeavey: The life and times of a New Zealand art dealer**
C. Bruce Ansley and Jane Ussher: Coast: A New Zealand journey
D. Damien Wilkins: Max Gate

15. Who hosted the 66th Emmy Awards?

A. Ellen DeGeneres
B. Tina Fey
C. **Seth Meyers**
D. Oprah

16. What cyclist was dubbed with the nickname “The Cannibal”?

A. Marco Pantani
B. Francesco Moser
C. Bernard Hinault
D. **Eddy Merckx**

17. What sport was Jack Dempsey famous for?

A. Baseball
B. Rugby
C. **Boxing**
D. Soccer

18. Which soccer team won the 2006 FIFA world cup in Germany?

A. **Italy**
B. Germany
C. Netherlands
D. France

19. In swimming, which stroke would one be performing when they put two arms over their head at once?

A. Front crawl
B. **Butterfly**
C. Backstroke
D. Trudgen

20. Who won the very first Grand Prix in motor racing in 1906?

A. Felice Nazzaro
B. Georges Boilot
C. **Ferenc Szisz**
D. Albert Clement

21. What is the sum of the dots on opposite sides of a standard die?

A. 5
B. **7**
C. 8
D. 9

22. Where were the 2004 Olympics?

A. **Athens, Greece**
B. Sydney, Australia
C. Beijing, China
D. Atlanta, United States

23. Which tennis player won 9 women's singles championships at Wimbledon?

A. Serena Williams
B. Martina Hingis
C. Anna Kournikova
D. **Martina Navratilova**

24. How many black squares are there on a chessboard?

A. 24
B. **32**
C. 40
D. 48

25. In which year was the water polo event introduced in Olympics?

A. 1904
B. 1912
C. **1900**
D. 1908

26. In Judo, the Black belt is the highest, what colour is the second highest?

A. Red
B. Purple
C. **Brown**
D. Green

27. Which of these horses won the Melbourne Cup three times in a row?

A. Americain
B. **Makybe Diva**
C. Sunline
D. Phar Lap

28. Who holds the record for the most victories in a row on the professional golf tour?

A. Jack Nicklaus
B. Arnold Palmer
C. **Byron Nelson**
D. Ben Hogan

29. The Wimbledon tennis tournament is played on what kind of surface?

A. Clay
B. **Grass**
C. Asphalt
D. Cement

30. Which Formula One driver claimed top honours at the British Grand Prix in Silverstone?

A. **Lewis Hamilton**
B. Fernando Alonso
C. Felipe Massa
D. Jenson Button

Appendix D. Questionnaire

Please state your gender.

How old are you?

What is your nationality?

How many siblings do you have?

If you are a student, what is your subject?

When you were 16 years of age, what was the income of your parents in comparison to other families in New Zealand?

How large was the community where you have lived the most time of your life?

How many people live in your household (including yourself)?

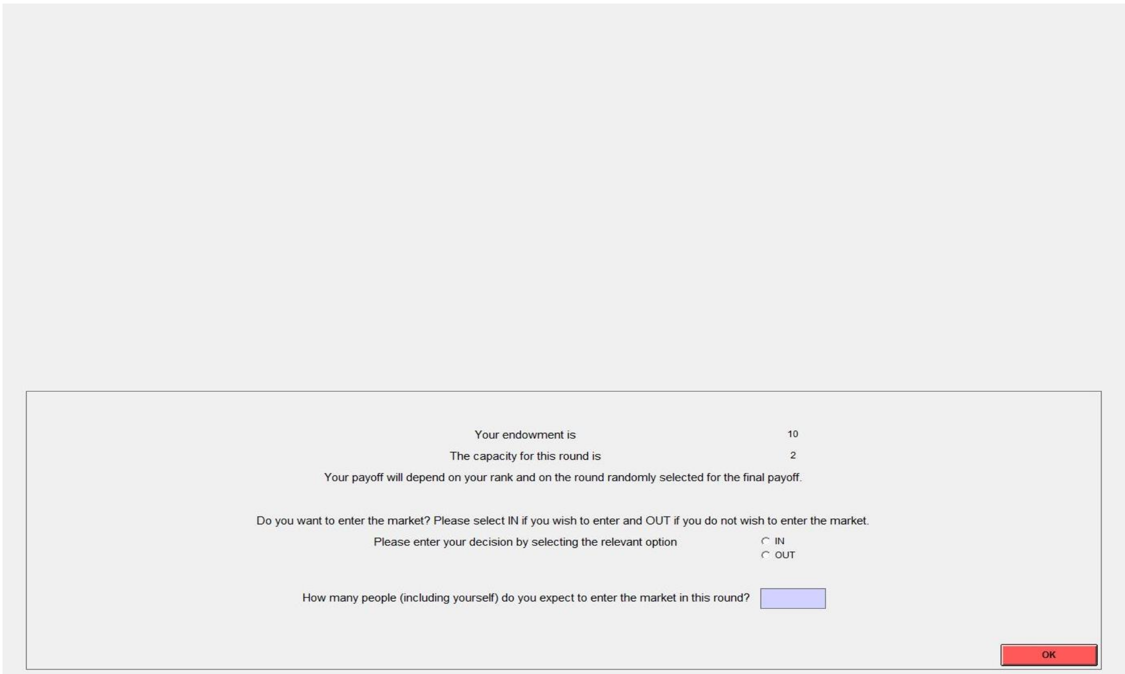
How large is your monthly budget (without expenses for accommodation)?

What share of your monthly expenses do you finance yourself?

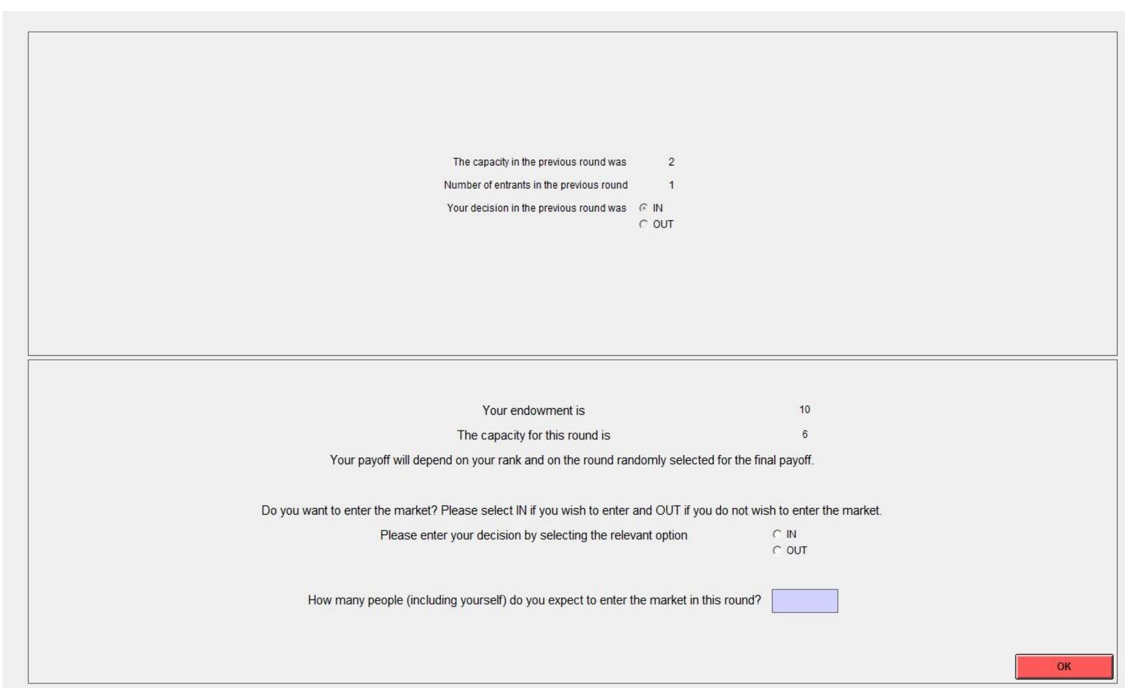
Can you state the percentage we can rely on the data you provided?

Appendix E. Screenshots From The Software

First round:



Second round (and all the following rounds with different “c”):



Appendix F. Human Ethics Committee Approval

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Lynda Griffioen
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2015/25

1 April 2015

Katarina Dankova
Department of Economics and Finance
UNIVERSITY OF CANTERBURY

Dear Katarina

The Human Ethics Committee advises that your research proposal “Overconfidence and excess entry” has been considered and approved.

Please note that this approval is subject to the incorporation of the amendments you have provided in your email of 30 March 2015.

Best wishes for your project.

Yours sincerely

Lindsey MacDonald
Chair
University of Canterbury Human Ethics Committee

University of Canterbury Private Bag 4800, Christchurch 8140, New Zealand. **www.canterbury.ac.nz**

F E S

References

Abbink, K., Irlenbusch, B., & Renner, E. (2000). The moonlighting game: An experimental study on reciprocity and retribution. *Journal of Economic Behavior & Organization, 42* (2), 265-277.

Abeler, J., Altmann, S., Kube, S., & Wibral, M. (2010). Gift Exchange and Workers' Fairness Concerns: When Equality Is Unfair. *Journal of the European Economic Association, 8* (6), 1299-1324.

Adams, J. S. (1965). Inequity in social exchange. *Advances in experimental social psychology, 2* (267-299).

Akerlof, G. A., & Yellen, J. L. (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics*, 255-283.

Alewell, D., & Nicklisch, A. (2006). Wage Differentials, Fairness, and Social Comparison: An Experimental Study of the Co-Employment of Permanent and Temporary Agency Workers. *MPI Collective Goods Preprint* (2006/8).

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49* (6), 1621.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of personality and social psychology, 68* (5), 804.

Anderson, C. M., & Putterman, L. (2006). Do non-strategic sanctions obey the law of demand? The demand for punishment in the voluntary contribution mechanism. *Games and Economic Behavior, 54* (1), 1-24.

Armantier, O. (2006). Do wealth differences affect fairness considerations? *International Economic Review, 47* (2), 391-429.

Arrow, K. J. (1982). Risk perception in psychology and economics. *Economic inquiry, 20* (1), 1-9.

Babcock, L., Wang, X., & Loewenstein, G. (1996). Choosing the wrong pond: Social comparisons in negotiations that reflect a self-serving bias. *The Quarterly Journal of Economics*, 1-19.

Baddeley, M. (2010). Herding, social influence and economic decision-making: socio-psychological and neuroscientific analyses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365* (1538), 281-290.

Barber, B. M., & Odean, T. (2000). Too many cooks spoil the profits: Investment club performance. *Financial Analysts Journal, 56* (1), 17-25.

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly journal of Economics*, 261-292.

Bardsley, N. (2010). *Experimental economics: Rethinking the rules*: Princeton University Press.

Ben-David, I., Graham, J. R., & Harvey, C. R. (2007). *Managerial overconfidence and corporate policies*: National Bureau of Economic Research.

Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *Quarterly journal of economics*, 871-915.

Bénabou, R., & Tirole, J. (2003). Self-knowledge and self-regulation: An economic approach. *The psychology of economic decisions, 1*, 137-167.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior, 10* (1), 122-142.

Bewley, T. F. (1999). *Why wages don't fall during a recession*: Harvard University Press.

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of personality and social psychology, 59* (5), 960.

Bolle, F. (1988). Learning to make good predictions in time series. In *Bounded Rational Behavior in Experimental Games and Markets* (pp. 37-50): Springer.

Bolton, G. E. (1991). A comparative model of bargaining: Theory and evidence. *American Economic Review*, 1096-1136.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 166-193.

Bosman, R., Sutter, M., & van Winden, F. (2005). The impact of real effort and emotions in the power-to-take game. *Journal of Economic Psychology, 26* (3), 407-429.

Bosman, R., & van Winden, F. (2002). Emotional hazard in a power-to-take experiment. *Economic Journal, 112* (476), 147-169.

Brandts, J., & Charness, G. (2011). The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics, 14* (3), 375-398.

Cabrales, A., & Charness, G. (2000). Optimal contracts, adverse selection, and social preferences: An experiment.

Camerer, C. (1992). The rationality of prices and volume in experimental markets. *Organizational Behavior and Human Decision Processes, 51* (2), 237-272.

Camerer, C., Kagel, H., & Roth, A. (1995). The handbook of experimental economics. *The Handbook of Experimental Economics*.

Camerer, C., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American economic review*, 306-318.

Camerer, C. F. (2003). Behavioural studies of strategic thinking in games. *Trends in cognitive sciences, 7* (5), 225-231.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty, 19* (1-3), 7-42.

Camerer, C. F., Loewenstein, G., & Prelec, D. (2004). Neuroeconomics: Why economics needs brains. *The Scandinavian Journal of Economics, 106* (3), 555-579.

Cárdenas, J. C., De Roux, N., Jaramillo, C. R., & Martinez, L. R. (2013). Is it my money or not? An experiment on risk aversion and the house-money effect. *Experimental Economics*, 1-14.

Carlsson, F., He, H., & Martinsson, P. (2013). Easy come, easy go: The role of windfall money in lab and field experiments. *Experimental Economics, 16* (2), 190-207.

Carpenter, J. P. (2007). The demand for punishment. *Journal of Economic Behavior & Organization, 62* (4), 522-542.

Caskey, J. (1985). Modeling the formation of price expectations: A Bayesian approach. *The American Economic Review*, 768-776.

Charness, G. (2004). Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics, 22* (3), 665-688.

Charness, G., Cobo-Reyes, R., Lacomba, J. A., Lagos, F., & Perez Sanchez, J. M. (2013). Social comparisons in wage delegation: Experimental evidence.

Charness, G., & Kuhn, P. (2007). Does pay inequality affect worker effort? Experimental evidence. *Journal of labor economics, 25* (4), 693-723.
Charness, G., & Kuhn, P. (2011). Lab labor: What can labor economists learn from the lab? *Handbook of Labor Economics, 4*, 229-330.
Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics, 117* (3), 817-869.
Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the dictator. *American Economic Review, 92* (4), 1218-1221.
Clark, A., Masclet, D., & Villeval, M. C. (2006). Effort and comparison income: Experimental and survey evidence.
Clark, A. E., Masclet, D., & Villeval, M. C. (2010). Effort and comparison income: Experimental and survey evidence. *Industrial & Labor Relations Review, 63* (3), 407-426.
Clark, A. E., & Senik, C. (2010). Who compares to whom? the anatomy of income comparisons in europe*. *The Economic Journal, 120* (544), 573-594.
Clark, J. (2002). House money effects in public good experiments. *Experimental Economics, 5* (3), 223-231.
Clark, K., & Sefton, M. (2001). The sequential prisoner's dilemma: evidence on reciprocation. *Economic Journal, 111* (468), 51-68.
Clemens, M. A. (2015). *The Meaning of Failed Replications: A Review and Proposal*: Institute for the Study of Labor (IZA).
Clingingsmith, D. (2013). Mental Accounts, Selective Attention, and the Mutability of Altruism: An Experiment with Online Workers. Working paper.
Costa-Gomes, M., & Zauner, K. G. (2001). Ultimatum bargaining behavior in Israel, Japan, Slovenia, and the United States: A social utility analysis. *Games and Economic Behavior, 34* (2), 238-269.
Cox, J. C., & Deck, C. A. (2005). On the nature of reciprocal motives. *Economic Inquiry, 43* (3), 623-635.
Cox, J. C., Friedman, D., & Gjerstad, S. (2007). A tractable model of reciprocity and fairness. *Games and Economic Behavior, 59* (1), 17-45.
Cox, J. C., Friedman, D., & Sadiraj, V. (2008). Revealed Altruism. *Econometrica, 76* (1), 31-69.
Cox, J. C., & Hall, D. T. (2010). Trust with private and common property: Effects of stronger property right entitlements. *Games, 1* (4), 527-550.
Cox, J. C., Sadiraj, K., & Sadiraj, V. (2008). Implications of trust, fear, and reciprocity for modeling economic behavior. *Experimental Economics, 11* (1), 1-24.
Cox, J. C., Servátka, M., & Vadovič, R. (2014). Status Quo Effects in Fairness Games: Reciprocal Responses to Acts of Commission vs. Acts of Omission. University of Canterbury working paper.
Croson, R., & Gächter, S. (2010). The science of experimental economics. *Journal of Economic Behavior & Organization, 73* (1), 122-131.
Croson, R. T. (1996). Information in ultimatum games: An experimental study. *Journal of Economic Behavior & Organization, 30* (2), 197-212.
Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under‐ and overreactions. *the Journal of Finance, 53* (6), 1839-1885.

Daniels, B. P., & Plott, C. R. (1988). Inflation and expectations in experimental markets. In *Bounded rational behavior in experimental games and markets* (pp. 198-218): Springer.

Danziger, L., & Katz, E. (1997). Wage secrecy as a social convention. *Economic Inquiry, 35* (1), 59-69.

Davis, D. D., & Holt, C. A. (1993). *Experimental economics*: Princeton university press.

De Bondt, W. F., & Thaler, R. H. (1995). Financial decision-making in markets and firms: A behavioral perspective. *Handbooks in Operations Research and Management Science, 9,* 385-410.

De Long, J. B., Summers, L. H., & Waldmann, R. J. (1991). The Survival of Noise Traders in Financial Markets. *Journal ofBusiness, 64* (1).

Deaves, R., Luders, E., & Luo, G. Y. (2009). An Experimental Test of the Impact of Overconfidence and Gender on Trading Activity. *Review of Finance, 13*, 555-575.

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social issues, 31* (3), 137-149.

Dewald, W. G., Thursby, J. G., & Anderson, R. G. (1986). Replication in empirical economics: The journal of money, credit and banking project. *The American Economic Review*, 587-603.

Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior, 47* (2), 268-298.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of personality and social psychology, 57* (6), 1082.

Duvendack, M., Palmer-Jones, R. W., & Reed, W. R. (2015). Replications in Economics: A Progress Report. *Scholarly Comments on Academic Economics*, *12* (2), 164-191.

Engelmann, D., & Strobel, M. (2000). *An experimental comparison of the fairness models by Bolton and Ockenfels and by Fehr and Schmidt*: Discussion Papers, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes.

Erev, I., & Rapoport, A. (1998). Coordination,"magic," and reinforcement learning in a market entry game. *Games and economic behavior*, *23* (2), 146-175.

Erkal, N., Gangadharan, L., & Nikiforakis, N. (2011). Relative Earnings and Giving in a Real-Effort Experiment. *American Economic Review, 101* (7), 3330-3348.

Falk, A., Fehr, E., & Fischbacher, U. (2003). On the nature of fair behavior. *Economic Inquiry, 41* (1), 20-26.

Falk, A., Fehr, E., & Fischbacher, U. (2008). Testing theories of fairness—Intentions matter. *Games and Economic Behavior, 62* (1), 287-303.

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior, 54* (2), 293-315.

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *science, 326* (5952), 535-538.

Falk, A., & Knell, M. (2004). Choosing the Joneses: Endogenous goals and reference standards. *The Scandinavian Journal of Economics, 106* (3), 417-435.

Fang, H., & Moscarini, G. (2005). Morale hazard. *Journal of Monetary Economics, 52* (4), 749-777.

Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review, 46* (4), 687-724.
Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review, 90* (4), 980-994.
Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica*, 833-860.
Fehr, E., & Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics, 114* (3), 817-868.
Festinger, L. (1954). A theory of social comparison processes. *Human relations, 7* (2), 117-140.
Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10* (2), 171-178.
Forbes, W. (2009). *Behavioural finance*: John Wiley & Sons.
Frank, R. H. (1984). Are workers paid their marginal products? *The American economic review*, 549-571.
Gächter, S. (2009). Improvements and future challenges for the research infrastructure in the field'experimental economics'.
Gächter, S., & Fehr, E. (2002). *Fairness in the labour market*: Springer.
Gächter, S., & Thöni, C. (2010). Social comparison and performance: Experimental evidence on the fair wage–effort hypothesis. *Journal of Economic Behavior & Organization, 76* (3), 531-543.
Garner, C. (1982). Experimental evidence on the rationality of intuitive forecasters. *Research in experimental economics, 2*, 113-128.
Gervais, S., & Goldstein, I. (2007). The positive effects of biased self-perceptions in firms. *Review of Finance, 11* (3), 453-496.
Gervais, S., Heaton, J. B., & Odean, T. (2011). Overconfidence, compensation contracts, and capital budgeting. *The Journal of Finance, 66* (5), 1735-1777.
Gervais, S., & Odean, T. (2001). Learning to be overconfident. *Review of Financial studies, 14* (1), 1-27.
Gneezy, U., & List, J. A. (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica, 74* (5), 1365-1384.
Golder, P. N., & Tellis, G. J. (1993). Pioneer advantage: Marketing logic or marketing legend? *journal of Marketing Research*, 158-170.
Greiner, B. (2004). An Online Recruitment System for Economic Experiments. In K. Kremer & V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen 2003* (Vol. 63, pp. 79-93). Göttingen: Ges. für Wiss. Datenverarbeitung.
Guala, F. (2005). *The methodology of experimental economics*: Cambridge University Press.
Güth, W., Huck, S., & Ockenfels, P. (1996). Two-level ultimatum bargaining with incomplete information: An experimental study. *The Economic Journal*, 593-604.
Güth, W., Königstein, M., Kovács, J., & Zala-Mezõ, E. (2001). Fairness within firms: The case of one principal and multiple agents. *Schmalenbach Business Review, 53* (2), 82-101.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3* (4), 367-388.

Hackbarth, D. (2008). Managerial traits and capital structure decisions. *Journal of Financial and Quantitative Analysis, 43* (04), 843-881.

Hamermesh, D. S. (2007). Viewpoint: Replication in economics. *Canadian Journal of Economics/Revue canadienne d'économique, 40* (3), 715-733.

Harrison, G. W. (2007). House money effects in public good experiments: Comment. *Experimental Economics, 10* (4), 429-437.

Heaton, J. B. (2002). Managerial optimism and corporate finance. *Financial management*, 33-45.

Henrich, J. (2000). Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *American Economic Review*, 973-979.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 73-78.

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science, 319* (5868), 1362-1367.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences, 24* (03), 383-403.

Ho, T.-H., & Su, X. (2009). Peer-induced fairness in games. *The American Economic Review, 99* (5), 2022-2049.

Hoelzl, E., & Rustichini, A. (2005). Overconfident: Do You Put Your Money On It?*. *The Economic Journal, 115* (503), 305-318.

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior, 7* (3), 346-380.

Hoffman, E., & Spitzer, M. L. (1985). Entitlements, rights, and fairness: An experimental examination of subjects' concepts of distributive justice. *Journal of Legal Studies, 14*, 259.

Hoorens, V., & Buunk, B. P. (1993). Social Comparison of Health Risks: Locus of Control, the Person ‐ Positivity Bias, and Unrealistic Optimism1. *Journal of Applied Social Psychology, 23* (4), 291-302.

Hunter, J. E. (2001). The desperate need for replications. *Journal of Consumer Research, 28* (1), 149-158.

Isaac, R. M., McCue, K. F., & Plott, C. R. (1985). Public goods provision in an experimental environment. *Journal of Public Economics, 26* (1), 51-74.

Jones, E. E. (1985). Major developments in social psychology during the past five decades. *The handbook of social psychology, 1*, 47-107.

Kagel, J. H., Kim, C., & Moser, D. (1996). Fairness in ultimatum games with asymmetric information and asymmetric payoffs. *Games and Economic Behavior, 13* (1), 100-110.

Kahneman, D. (1988). Experimental economics: A psychological perspective. In *Bounded rational behavior in experimental games and markets* (pp. 11-18): Springer.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American economic review*, 728-741.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of political Economy*, 1325-1348.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American psychologist, 39* (4), 341.

Klayman, J., Soll, J. B., Gonzalez-Vallejo, C., & Barlas, S. (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational behavior and human decision processes, 79* (3), 216-247.

Knez, M. J., & Camerer, C. F. (1995). Outside options and social comparison in three-player ultimatum game experiments. *Games and Economic Behavior, 10* (1), 65-94.

Konow, J. (1996). A positive theory of economic fairness. *Journal of Economic Behavior & Organization, 31* (1), 13-35.

Krueger, A. B., & Summers, L. H. (1988). Efficiency wages and the inter-industry wage structure. *Econometrica: Journal of the Econometric Society*, 259-293.

Kruger, J. (1999). Lake Wobegon be gone! The" below-average effect" and the egocentric nature of comparative ability judgments. *Journal of personality and social psychology, 77* (2), 221.

Kube, S., Maréchal, M., & Puppe, C. (2012). The currency of reciprocity-gift-exchange in the workplace. *American Economic Review, 102* (4).

Kuhn, T. S. (1962). The Structure of Scientific Revolutions Vol.

Lawler, E. E. (1990). *Strategic pay: Aligning organizational strategies and pay systems*: Jossey-Bass.

Leventhal, G. S. (1976a). The distribution of rewards and resources in groups and organizations. *Advances in experimental social psychology, 9*, 91-131.

Leventhal, G. S. (1976b). *Fairness in social relationships*: General Learning Press Morristown, NJ.

Lewis, K. K. (1989). Can learning affect exchange-rate behavior?: The case of the dollar in the early 1980's. *Journal of Monetary Economics, 23* (1), 79-100.

Libby, R., & Lipe, M. G. (1992). Incentives, effort, and the cognitive processes involved in accounting-related judgments. *Journal of Accounting Research*, 249-273.

Lundeberg, M. A., Fox, P. W., & Puncochar, J. (1994). Highly confident but wrong: Gender differences and similarities in confidence judgments. *Journal of educational psychology*, *86* (1), 114.

Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *The journal of finance, 60* (6), 2661-2700.

Marwell, G., & Ames, R. E. (1981). Economists free ride, does anyone else?: Experiments on the provision of public goods, IV. *Journal of public economics, 15* (3), 295-310.

Menkhoff, L., Schmidt, U., & Brozynski, T. (2006). The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence. *European Economic Review, 50* (7), 1753-1766.

Messick, D. M., Bloom, S., Boldizar, J. P., & Samuelson, C. D. (1985). Why we are fairer than others. *Journal of Experimental Social Psychology, 21* (5), 480-500.

Miller, A. G., Ashton, W. A., McHoskey, J. W., & Gimbel, J. (1990). What price attractiveness? Stereotype and risk factors in suntanning behavior. *Journal of Applied Social Psychology, 20* (15), 1272-1300.

Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior and Human Decision Processes, 103* (2), 197-213.

Nerlove, M. (1958). Adaptive expectations and cobweb phenomena. *The Quarterly Journal of Economics*, 227-240.

Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics, 92* (1), 91-112.

Nikiforakis, N., & Engelmann, D. (2011). Altruistic punishment and the threat of feuds. *Journal of Economic Behavior & Organization, 78* (3), 319-332.

Nikiforakis, N., & Normann, H.-T. (2008). A comparative statics analysis of punishment in public-good experiments. *Experimental Economics, 11* (4), 358-369.

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of finance, 53* (5), 1775-1798.

Odean, T. (1999). Do Investors Trade Too Much? *American Economic Review*, 1279-1298.

Ostrom, E., & Walker, J. (2005). *Trust and reciprocity: Interdisciplinary lessons from experimental research* (Vol. 6): Russell Sage Foundation Publications.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review*, 404-417.

Oxoby, R. J., & Spraggon, J. (2008). Mine and yours: Property rights in dictator games. *Journal of Economic Behavior & Organization, 65* (3), 703-713.

Pereira, P. T., & Silva, N. (2006). Positive and negative reciprocity in the labor market. *Journal of Economic Behavior & Organization, 59* (3), 406-422.

Perloff, L. S., & Fetzer, B. K. (1986). Self–other judgments and perceived vulnerability to victimization. *Journal of Personality and social Psychology, 50* (3), 502.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 1281-1302.

Reinstein, D., & Riener, G. (2012). Decomposing desert and tangibility effects in a charitable giving experiment. *Experimental Economics, 15* (1), 229-240.

Requate, T., Waichman, I., & Siang, C. (2011). On the role of social wage comparisons in gift-exchange experiments. *Economics Letters, 112* (1), 75-78.

Roll, R. (1986). The hubris hypothesis of corporate takeovers. *Journal of business*, 197-216.

Roth, A. E. (1994). Lets keep the con out of experimental econ.: a methodological note. In *Experimental Economics* (pp. 99-109): Springer.

Roth, E. A. (2001). Form and function in experimental design. *Behavioral and Brain Sciences, 24* (3), 427-428.

Rutström, E. E., & Williams, M. B. (2000). Entitlements and fairness:: an experimental study of distributive preferences. *Journal of Economic Behavior & Organization, 43* (1), 75-89.
Schmalensee, R. (1976). An experimental study of expectation formation. *Econometrica: journal of the Econometric Society*, 17-41.
Scott, R. E. (2003). A theory of self-enforcing indefinite agreements. *Columbia Law Review*, 1641-1699.
Selten, R. (1967). Ein Marktexperiment. W: H. Sauermann (red.). *Beiträge zur experimentellen Wirtschaftsforschung, 1*.
Shefrin, H. (2001). Behavioral corporate finance. *Journal of Applied Corporate Finance, 14* (3), 113-126.
Slonim, R., & Roth, A. E. (1998). Learning in high stakes ultimatum games: An experiment in the Slovak Republic. *Econometrica*, 569-596.
Smith, V. L., & Walker, J. M. (1993). Monetary rewards and decision cost in experimental economics. *Economic Inquiry, 31* (2), 245-261.
Soll, J. B. (1996). Determinants of overconfidence and miscalibration: The roles of random error and ecological structure. *Organizational Behavior and Human Decision Processes*, *65* (2), 117-137.
Straub, P. G., & Murnighan, J. K. (1995). An experimental investigation of ultimatum games: Information, fairness, expectations, and lowest acceptable offers. *Journal of Economic Behavior & Organization, 27* (3), 345-364.
Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica, 47* (2), 143-148.
Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological bulletin*, *103* (2), 193.
Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science, 4* (3), 199-214.
Thaler, R., & Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science, 36* (6), 643-660.
Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, S251-S278.
Weinberg, B. A. (2009). A model of overconfidence. *Pacific Economic Review*, *14* (4), 502-515.
Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology*, *39* (5), 806.
Weinstein, N. D. (1982). Unrealistic optimism about susceptibility to health problems. *Journal of behavioral medicine, 5* (4), 441-460.
Weinstein, N. D. (1984). Why it won't happen to me: perceptions of risk factors and susceptibility. *Health psychology, 3* (5), 431.
Weinstein, N. D. (1987). Unrealistic optimism about susceptibility to health problems: Conclusions from a community-wide sample. *Journal of behavioral medicine*, *10* (5), 481-500.
Weinstein, N. D. (1989). Optimistic biases about personal risks. *Science*, *246* (4935), 1232-1233.

Weinstein, N. D. (1998). Accuracy of smokers' risk perceptions. *Annals of behavioral medicine, 20* (2), 135-140.

Weinstein, N. D., & Klein, W. M. (1995). Resistance of personal risk perceptions to debiasing interventions. *Health psychology, 14* (2), 132.

Wilson, R. K., & Isaac, R. M. (2007). Political Economy and Experiments. In *The Political Economist: Newsletter of the Section on Political Economy*.

Windschitl, P. D., & Chambers, J. R. (2004). The dud-alternative effect in likelihood judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30* (1), 198.