Exploring Sequential Choice Task Strategies

by

Jesse Langstaff

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2011

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

_____

Jesse Langstaff

Abstract

The current study provides evidence that individuals tend to adopt an integrative choice strategy when making sequential decisions under conditions of uncertainty. This contrasts with prior literature which proposes that decisions are made one at a time in isolation from one another (Camerer et al., 1997). By creating an experimental work task where only wage quality and feedback are manipulated, the resulting changes in intertemporal substitution between work and leisure are observed. In Experiments 1-3, a positive relationship between wages and time spent working that did not depend on task experience was observed. These results suggest that decisions are being made in consideration of other decisions, as isolated decisions would yield a negative relationship between wages and time spent working. In Experiment 4 these results were mitigated by the difficulty in differentiating between low and high wage quality days. These findings are taken to suggest that the results of prior studies are primarily due to self-control issues that subjects faced, which are not present in the present study.

Acknowledgments

I would like to extend my appreciation to Derek Koehler for his assistance in this project specifically, and my development as a researcher in general. I would also like to thank Britt Anderson and Jonathan Fugelsang who provided helpful criticism in the development of this document. Finally, I would like to thank Adriel Orena for assisting with data collection.

Table of Contents

List of Figures

List of Tables

Introduction

When faced with making sequential decisions under conditions of uncertainty, people often fail to integrate their decisions with ones made prior to the current decision point as well as without consideration to future time points in which subsequent decisions will be made. This process of making decisions one at a time, in isolation from one another, results in suboptimal performance as individuals fail to adopt a broader, more integrative choice strategy that incorporates beneficial tradeoffs across the decision sequence.

For instance, Camerer et al. (1997), in their observations of the working behaviour of New York City cab drivers, found that drivers tended to work longer hours on low-wage than on high-wage days. These cab drivers owned, rented or drove fleet cabs and were free to determine how long they spent driving. The wages that the drivers earned often fluctuated across days, but was usually consistent within a day; that is when a day was busy in the morning and drivers could spend less time searching for fares, they could expect that the latter half of the day would also be busy.

By working more on low-wage than on high-wage days, drivers displayed a negative elasticity of labour supply that was significantly different from zero and in inexperienced drivers close to negative one. This meant that on average when drivers faced a day in which wages increased by five percent, it would be expected that they would supply five percent less labour; this would be done by spending less time working (i.e., quitting earlier). As cab drivers became more experienced, Camerer et al. observed this negative elasticity shrink towards zero and hypothesised that this was due to drivers either learning a more efficient strategy to allocate their time, or due to self selection out of the job market by cab drivers with negative elasticities.

There are many possible explanations that could account for the observed negative elasticities of these cab drivers that Camerer et al. dismissed and which will be briefly summarized before discussing Camerer er al.'s preferred explanation. First, it could have been that drivers were liquidity constrained (i.e., short on cash and unable to borrow), and had to work long hours on low wage days in order to earn enough income to pay for their expenses. However, since even drivers who bought their own licence, which at the time cost approximately $130,000, displayed negative elasticities this reason could not explain their behaviour. Alternatively, since what made a high-wage day have high wages was driving more customers and spending less time idly searching, cab drivers could have been more fatigued and tired sooner than on a more relaxed low wage-day and as a result quit the day early to rest. However, drivers self-reported that the opposite was true: low-wage days which contained more searching for customers than driving customers were felt as being more tiring. Finally, Camerer et al. were able to dismiss the possibility that drivers simply drove to obtain happy endings or take a lot of breaks on slow days, both of which would cause working hours on low-wage days to be longer, by controlling for these possibilities when analyzing their dataset.

Having rejected alternative causes of drivers negative elasticities, Camerer et al. proposed the idea that cab drivers simply drove each day until they reached a target income. By performing this simple strategy, drivers would need to drive longer on low-wage days as it would take longer to reach their target; whereas on high-wage days the target would be reached in comparatively less time and as result they would not need to drive as many hours as on a low-wage day. However, this proposed targeting behaviour of cab drivers is inefficient, as drivers could obtain more income and subsequently more leisure time by driving more on

high-wage and less on low-wage days and then substituting income and leisure across these types of days. Such a strategy would result in drivers having positive wage elasticities, opposite to what Camerer et al. observed.

Economic Theory of Labour Supply

This optimal strategy, of working more on high-wage and less on low-wage days, that yields positive wage elasticities, is how economic theories of labour supply predict workers will behave. In these models, workers attempt to maximize their total lifetime utility, where utility represents a measure of happiness or satisfaction that can be decomposed into contributions from leisure and consumption, while being constrained in how much time and income they have. For example, while someone would prefer to consume an infinite amount of chocolate cake while spending all their time in leisure, the amount of chocolate cake they can buy is restricted to the amount they can afford and the amount of leisure time they have is restricted by the total amount of time in the period of interest, less time spent working in order to earn the income to afford the chocolate cake. This theory can be expressed mathematically (Farber, 2005), with the utility from consumption and leisure for a worker on day t being additively separable within the period

$$U_t = a(c_t) + b(l_t) \tag{1}$$

where *c* represents consumption of goods and *l* is daily leisure consumption. The lifetime utility function, which is additively separable across time, defined over a set of *T* periods is

$$U = \sum_{t=0}^{T} (1+\rho)^{-t}[a(c_t) + b(l_t)] \tag{2}$$

where ρ is the temporal discount rate, and $a(\cdot)$ and $b(\cdot)$ have positive first derivatives and negative second derivatives. That is both $a(\cdot)$ and $b(\cdot)$ are subject to diminishing returns. The lifetime budget constraint which determines the highest accessible utility curve is described by

$$Y_0 + \sum_{t=0}^{T}(1+r)^{-t} y_t(1-l_t) = \sum_{t=0}^{T}(1+r)^{-t} c_t \tag{3}$$

where the price of consumption goods is normalized to one, r is the discount rate, Y_0 represents initial wealth , $1 - l_t$ represents work hours and $y_t(\cdot)$ represents daily earnings. The Lagrangian expression for constrained maximization of the lifetime utility function is

$$V = \sum_{t=0}^{T}(1+\rho)^{-t}[a(c_t) + b(l_t)] + \lambda\left\{Y_0 + \sum_{t=0}^{T}(1+r)^{-t}[y_t(1-l_t) - c_t]\right\} \tag{4}$$

where is interpreted as the marginal utility of lifetime wealth. This expression is maximized with respect to c_t and l_t and the first order conditions are

$$\frac{\partial V}{\partial c_t} = a(c_t) - \lambda\theta^t = 0 \tag{5}$$

$$\frac{\partial V}{\partial l_t} = b(l_t) - \lambda\theta^t y_t(1-l_t) = 0 \tag{6}$$

$$\frac{\partial V}{\partial \lambda} = Y_0 + \sum_{t=0}^{T}(1+r)^{-t}[y_t(1-l_t) - c_t] = 0 \tag{7}$$

where $\theta = (1+\rho)/(1+r)$. Solving equations (5) and (6) for $\lambda\theta^t$ yields the result

$$y_t(1-l_t) = \frac{b(l_t)}{a(c_t)} \tag{8}$$

Equation (8) implies that workers should select hours such that the marginal wage from working an additional increment of time is equal to the marginal rate of substitution of leisure for goods within a single period. This dynamic model of labour supply predicts that

temporary changes in wages should be followed by a simultaneous change in labour supply. This is because a temporary unexpected wage change does not affect lifetime wealth, but results in an immediate change in the price of leisure (i.e., its opportunity cost in the form of foregone wage).

Myopic Loss Aversion

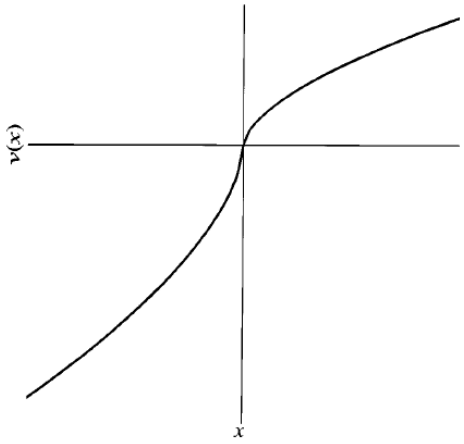
In contrast to the prediction from labour economics theory that workers perform a constrained optimization calculation to determine the optimal allocation of leisure time and consumption, which would always yield positive elasticities, Camerer et al.'s income targeting strategy is proposed to be a consequence of myopic loss aversion. Myopia refers to the fact that drivers are short-sighted and only consider the current day when making their working decisions, and loss aversion refers to the fact that drivers experience the pain of losses more than the happiness derived from gains of equivalent magnitude. These two ideas, and relevant research, are further discussed below.

Both myopia and loss aversion are necessary components of drivers' income targeting strategy, as without either drivers would tend to adopt a more beneficial strategy. If drivers were not myopic, they would be able to consider their decisions across a larger time horizon (rather than setting a fixed daily target) and be able to make tradeoffs between days, working more on some days in order to pay for days in which they took more leisure time. Even the smallest amount of planning would yield positive wage elasticities, as even considering only a two-day horizon would lead to some form of substitution. As is elaborated below, this account also requires that drivers exhibit loss aversion relative to a reference point determined by the daily target; that is, it must be more painful to fall just short of the target than it is pleasurable to exceed it by the same amount. Without the assumption of loss aversion, drivers

would not necessarily be expected to stop driving when and only when they reach their daily income target.

The concept of goals, such as a cab driver's daily income target, acting as a reference point and being subject to loss aversion was mentioned by Tversky and Kahneman (1979), who discussed that "level of aspiration" rather than the status quo could serve as a reference point. Heath et al. (1999) later elaborated on this idea and experimentally tested it. They found that by using Prospect Theory's (Tversky & Kahneman, 1979) value function (Figure 1) it is possible to explain how motivation towards pursuing a goal alters in relationship to the distance from the goal, or whether or not the goal has already been completed, or is just about to be completed.

Figure 1. Prospect Theory Value Function



For example, a person's goal acts as a reference point in which either success or failure can be determined, and where success yields positive emotions and failure results in negative emotions to be experienced. This is described by the reference point of the Prospect Theory value function, located of at the origin of the graph it divides areas of losses (left side of y-axis) from areas of gains (right side of y-axis). If someone has fallen short of a goal they set,

even if their performance is better than any prior instance, they will experience negative emotions.

These negative feelings, incurred when below a goal, are experienced as being more painful than a similar position above the goal would be pleasurable, as implied by loss aversion where the value function has a steeper slope in the domain of losses than in the domain of gains. Therefore, as a result of being below a goal and facing the prospect of not reaching their goal, individuals will exert more effort to increase their performance to bring them one unit closer to the goal than someone who has already achieved the goal and whose next unit of effort only serves to move them farther past the already obtained goal.

A final feature of the value function which plays an important role in explaining goal pursuit behaviour is diminishing sensitivity, which results in the value function being concave for the region of gains and convex for the area of losses. This has two important implications. First, it implies that individuals will exert less effort after they have completed a goal and each further increment moves them farther from their original goal. Conversely, more effort should be exerted when below a goal and approaching it as each increment towards the goal results in large jumps in value. That is, when approaching the goal, each step brings more value than the one before it; by contrast, each step past the goal brings less value than the one before it. Second, the convex nature of the value function when in the domain of losses implies that individuals will be more risk seeking in their actions, while those who are on the concave portion in the domain of gains will be more risk averse.

Heath et al.'s (1999) analysis offers a simple explanation of results from prior goal studies showing how goals can increase task perseverance, something that the New York City

cab drivers appeared to be doing by working longer on low wage days until they reached a target income.

Larrick et al. (2009) further demonstrated the strong effect goals can have in shaping our behaviour, even if that behaviour ends up being detrimental. In a task that was to be repeated, two participants were paired together and asked to divide $8.00 between each other, with one participant proposing a distribution and the other either accepting or rejecting it. Rejection resulted in neither participant receiving any of the money, thus asking for a lopsided distribution was considered a risky strategy. However, participants who had previously set a target goal for themselves, rather than simply trying to do their best, often made more risky demands and even held to them over multiple rounds, resulting in more occurrences of rejections. This attitude of taking greater risks, was replicated in an individual task where participants could choose from a number of different gambles. Some participants were asked to set a goal for how much they were trying to win, while others were simply told to do their best. Similar to the negotiation task, those in the goal conditions pursued riskier gambles and even neglected to choose the gamble which would have given them a guaranteed value.

Along with loss aversion, Camerer et al (1997) also proposed myopia as a critical component of driver's targeting behaviour, reflecting the fact that drivers neglected to consider future work days when deciding how much they were willing to drive on the current day. This one day at a time mentality, results in a more risk averse style of decision making than one that would involve considering a broader set of days, such as a week or month. The increase in risk aversion as a result of making decisions one at a time, as opposed to all at once, was discussed by Gneezy and Potters (1997), who presented participants with a sequence of twelve identical and independent rounds of a lottery and an endowment which

participants drew from to put towards how much they wanted to place on each gamble. However, one group of participants played the gambles one by one and another group played the gambles in blocks of three. This manipulation of valuation period resulted in participants who played the independent gambles overall betting less than those who played the grouped gambles.

To better understand how evaluating one gamble at a time results in more risk averse behaviour than three simultaneous plays of the same gamble, consider an individual who is loss averse and as a result weighs losses more than gains. The rate at which they do this can be described as $\gamma > 1$. If the probability of losing \$1 on an individual lottery is 2/3 and the probability to win \$2.50 is 1/3 the expected utility of the single lottery can be described as

$$EU = \frac{2}{3}\gamma(-1) + \frac{1}{3}(2.50) \tag{9}$$

and only results in a positive expected utility if the rate at which losses are weighed to gains (γ) is less than 1.25. In contrast, if the three lotteries are evaluated at the same time the expected utility would be expressed as

$$EU = \frac{8}{27}\gamma(-3) + \frac{1}{27}(7.50) + \frac{6}{27}(4.00) + \frac{12}{27}(0.50) \tag{10}$$

and only results in a positive expected utility if the rate at which losses are weighed to gains (γ) is less than 1.56 since by evaluating the probability of a loss now drops to 30% from 67%.

Myopic loss aversion has been shown to result in suboptimal behaviour in other areas of decision making, indicating that it may be a generalizable bias affecting a wide breadth of decision making problems. Benartzi and Thaler (1995) demonstrated how myopic loss aversion could be used to explain the equity premium puzzle, an anomaly in which despite stocks having significantly greater returns than bonds, investors still choose to hold bonds and

require a premium to hold equity. Typical models of risk aversion could not account for the premium placed on equity, as models predicted that investors would need relative risk aversion coefficients of 30 or more to justify the premium (Mehra & Prescott, 1985). However, the premium can be explained using myopic loss aversion by considering the period of time in which investors evaluate their portfolio. Similar to the gambling problem presented above, those who evaluate their portfolio over a narrow time horizon, will be more loss averse than investors who evaluate over a broad time horizon. Thus the larger premium is needed to offset the loss aversion induced by shorter (narrow) evaluation periods.

The effects of manipulating the evaluation period on investors performance has also been observed more directly in the field at the Tel Aviv stock exchange, which contains securities that shift from daily to weekly trading. Kliger and Levit (2009) noted that myopic loss aversion predicts a negative relation between risk aversion and the length of evaluation period, that the longer the evaluation period the lower the expected return is and that under longer evaluation periods there would be a reduced sensitivity to economic events. Examining the results of stocks that changed from daily to weekly trading, they found that these predictions did in fact hold: stocks subject to weekly trading had lower returns and fluctuated less as a result of economic events than when traded daily.

While the previous two studies emphasize the effects of evaluation periods on decisions, it is also necessary to understand the role of experience as it influences myopic loss aversion. Camerer et al. (1997) proposed that drivers' elasticities became less negative with more experience due to drivers adopting a strategy other than targeting which they arrived at through learning from experience on the job; thus myopic loss aversion should be less prevalent among experienced drivers when compared to novice cab drivers. However, two

similar studies (Haigh & List, 2005 and Eriksen & Kvaloy, 2010) found that when asked to perform a task similar to the one previously outlined by Gneezy and Potters (1997), experienced stock traders and those involved in financial markets actually exhibited greater amounts of myopic loss aversion than students who participated in the same task. That is, despite their experience, stock traders were even more prone than non-experts to gamble more when the evaluation period was longer.

All of these examples of myopic loss aversion have mostly involved financial markets, or those who work in them, but myopia occurs outside of this area as well. For example, Read and Loewenstein (1995) observed that when asking participants once a week over the course of three weeks what snacks they would like, participants frequently selected snacks that they had chosen during a prior week. In contrast, participants who made a nonbinding choice of three weeks' worth of snacks all at once, tended to diversify their selection more. In other words, when thinking about a broad set of options, instead of a narrower set, participants tended to diversify their selections more. This implies that simply changing how a question is framed (e.g., narrow or broadly), can influence whether decisions are made in isolation from one another or not. This diversification bias was also replicated during Halloween. Children visiting two neighboring houses were either asked to choose two candy bars from one house or one candy bar each from two houses, both offering two types of candy bars. Children presented with the two choices at once opted to take different chocolate bars, while those choosing one from each house predominately selected the same one.

The Current Experiments

This prior literature demonstrates that myopic loss aversion can explain results which diverge from the predictions of standard economic theory. It therefore seems plausible then

that myopic loss aversion could also underlie driver's suboptimal allocation of working hours across days, as proposed by Camerer et al. (1997), which contradicts what theories of labour supply would predict. However, Camerer et al.'s findings and explanation are all purely correlational and do not truly establish a causal link between wages and hours worked. Due to the uncontrolled nature of the field study, it is possible that something other than wage quality could be what is affecting the amount of time drivers choose to work. For example, it could be possible that high-wage days are also days in which cab drivers would benefit the most from leisure and therefore simply quit earlier because they do not want to be at work on those days. A simple example of this would be in a University town, the busiest time for cab drivers would be evenings and weekends (particularly weekend evenings) when students are most likely to spend time out. However, these times are also ones in which the driver could spend with friends and family. As a result of the possibility of some confounding variable, it is only truly possible to understand the causal relationship between wages and hours worked by manipulating wages and observing how participants respond to the change via a shift in work hours while holding all other factors constant.

Along a similar vein, Camerer et al.'s measure of experience's role on working behaviour is susceptible to issues with confounding variables, which is in turn complicated by the fact that it varied between participants and not within. Therefore, it is unclear whether or not the apparent behavioural differences are actually a result of experience or something that is simply inherently different between the two distinct populations of cab drivers. However, in a controlled experimental setting the role of experience can be observed longitudinally within subjects, thus allowing us to see how an individual's performance actually changes across time (i.e., with experience).

Along with these issues, evidence in support of income targeting is not made very clear, with the exception of the occurrence of negative wage elastiticities and some explicit questionnaire responses by a small sample of fleet managers. By creating an experimental setting that attempts to replicate cab driver's environment, it becomes possible to also perform manipulations that attempt to impact a component of income targeting, such as myopic loss aversion. As will be discussed shortly, the following experiments seek to manipulate whether participants behave in a myopic fashion or not, to determine whether myopia is indeed an underlying mechanism that contributes to negative wage elasticities.

Due to these shortcomings with prior correlational studies, the present paper attempts to better understand the wage – hours link by exploring the relationship experimentally rather than correlationally. In addition, attention will be paid to the moderating effect of experience on task performance, examining whether or not a similar trend with greater experience towards positive elasticities exists as observed by Camerer et al. In order to test Camerer et al.'s account, a simulated work task was designed in which participants completed a sequence of work days in which they could choose either to work and acquire utility through income consumption or take leisure time and acquire utility through leisure consumption. Both income and leisure were converted, with diminishing returns, into a satisfaction (i.e., utility) score, which participants were tasked with maximizing over the course of the experiment.

In order to observe how work hours fluctuated with respect to changes in wages, days of varying wage quality were randomly distributed over the course of the experiment. Participants' decision about how much and when to work directly determined the satisfaction scores they received, which in turn determined an actual monetary payment they received for participating in the experiment. That is participants who worked more on high wage days and

less on low wage days would accumulate more satisfaction points over the course of the experiment than those who worked less on high wage days and more on low wage days, and since participants received a cash reward based on satisfaction points it was in their best interest to distribute their work and leisure time in an optimal fashion.

With this motivating factor, we anticipated seeing a gradual shift in strategies over the course of the experiment, reminiscent of what Camerer et al. observed. Initially, participants might begin by using simple strategies, such as income targeting, and through some exploration from that point learn and move to more optimal strategies such as intertemporal substitution. Thus we would be better able to identify experience's role in task performance, in contrast to Camerer et al.'s observation which could in fact have been caused by cab driver's performing poor strategies and then quitting due to poor returns, leaving only drivers who had always driven optimally to continue driving.

In addition we attempted to test whether myopic loss aversion was an underlying cause of income targeting by manipulating the framing of task performance feedback, such that some participants received narrowly framed (i.e., more frequent, e.g., "daily") feedback and others received broadly framed (i.e., less frequent, e.g., "weekly") feedback. Should myopic loss aversion be an underlying cause of income targeting, it would then follow that those participants in the narrow feedback condition would show more signs of income targeting (such as negative elasticities) than those in the broad feedback condition. However, it is also possible that the opposite may in fact occur; i.e., those in the broad feedback condition may display more income targeting behaviour because they receive less frequent feedback by which to gauge or improve performance, whereas those in the narrow feedback condition can more readily experiment and see the results of their different allocations of work and leisure.

This could also play a large role in the time it takes to shift to a new strategy as a result of experience, with participants receiving narrow feedback being able to adopt the optimal strategy sooner than those receiving broad feedback.

Given the findings of Camerer et al. and the proposed experimental manipulations, this study aims to test the following three hypotheses.

The primary finding of Camerer et al. was the negative correlation between hours worked and hourly wages. This finding was proposed to be the result of income targeting. By experimentally manipulating wages, while holding other factors constant, it will be possible to see if this negative relationship still occurs. Without the occurrence of negative wages, there would be no purpose to posit income targeting as an explanation for people's behaviour. It follows that:

H1: Participants will show a negative relationship between hours worked and hourly wage.

The negative relationship between hours worked and wages observed by Camerer et al. attenuated as cab drivers gained more experience. While this was thought to be because drivers initially started with an inferior strategy such as income targeting and then through experimentation moved into intertemporal substitution, this in fact may not be the case as already discussed. Nonetheless, it is still expected that experience and practice with the simulated task will have positive benefits towards participants' performance. No matter what strategy participants initially begin with, it is unlikely that it will be the most optimal, and also reasonable to assume that they will not get any worse as they have a financial incentive to perform better. If there is in fact no effect of experience in moderating working behaviour, then it becomes possible that what Camerer et al. observed was simply a difference between

two different groups, specifically the experienced drivers could have been always performing intertemporal substitution while the novice drivers who targeted simply did not remain as cab drivers very long. It follows that:

H2: The relationship between hours worked and wages will be positively moderated by the role of experience.

A fundamental component of income targeting is myopic loss aversion. Thus it would be expected, that by manipulating whether people view the problem of accruing satisfaction points as something to be taken one day at a time, versus being the result of the consideration of multiple days in a period, will influence what strategy participants use. Those who receive narrow feedback will have the actions of the current day made more salient and therefore should approach decisions in a myopic fashion, which should lead to the use of income targeting (or at least should not decrease its use if targeting is the default or spontaneously adopted strategy). In contrast, those participants who receive broad feedback about a set of days should be more likely to consider the tradeoffs that can be made across those days. It follows that:

H3: Relative to those given narrow feedback, participants given broad feedback will move away from a targeting strategy toward intertemporal substitution more rapidly as a function of experience.

Experiment 1

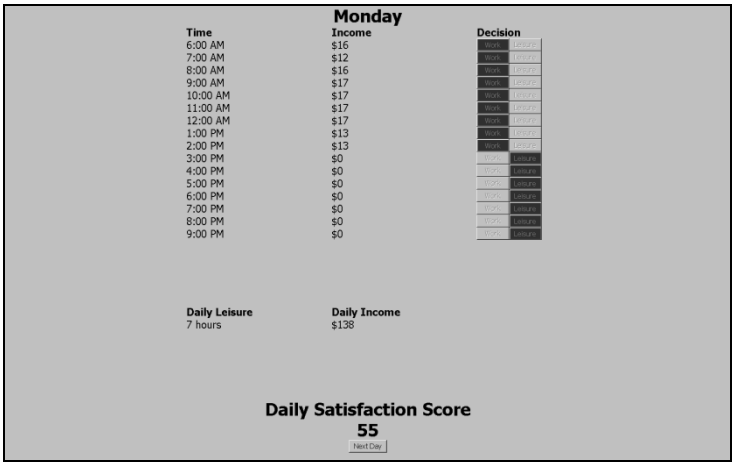
Methods

Participants. Participants were 60 undergraduate students (41 female), recruited from introductory psychology courses, who were told that they could receive between $2-$6 for their participation depending on their performance in addition to course credit.

Apparatus. Each participant completed the experiment on a personal computer. All stimuli and instructions were displayed on the monitor, and all responses were given by using the mouse to select options presented onscreen. A research assistant provided a summary of the instructions and ensured that participants understood them.

Procedure. The experiment consisted of 100 trials, each consisting of 16 work/leisure decisions. Participants were asked to imagine themselves in the role of a cab driver who is trying to have a satisfactory lifestyle. To add to this immersion, trials were known as days and the 16 work/leisure decisions occurred (one per "hour") within the day. A daily cab rental fee of $50 per day was deducted from income.

Figure 2. Example of a completed day with daily feedback



Over the course of a day participants made a decision each hour as to whether they would work, and earn a wage which accumulated as income, or take leisure, which

accumulated as leisure hours. Both income and leisure hours were displayed to participants and updated to reflect their decisions. When a new day began these values were reset to zero. Participants chose to spend an hour doing work or leisure by clicking the button titled "Work" or the button titled "Leisure" which were displayed next to the current hour. If work was chosen the wage for that hour then appeared next to the hour where previously a "?" existed. However, once the participant elected to take leisure, no wage for that hour was displayed; furthermore, the participant was then only able to make selections of leisure for the rest of the day. Because only decisions regarding how long to work on any given day were of interest, removing decisions about when to work during the day (e.g., mornings vs. afternoons) served to remove the possibility of participants attributing their outcomes to their within-day working patterns.

At the end of every day income and leisure hours were combined to calculate a daily satisfaction score (a proxy for economic utility) according to the following equation:

$$daily\ satisfaction\ score = (25Leisure)^{.88} + Income^{.88} - 115$$

This daily satisfaction score equation specifies an additive relationship between income and leisure in determining daily satisfaction, both of which were subject to diminishing returns as is commonly assumed in models of economic utility. The rate at which the contribution of either leisure or income decreased was determined by the exponent .88, which is also the estimated value of the exponent parameter used in forming cumulative prospect theory's value function (Tversky & Kahneman, 1992). Before experiencing diminishing returns the value of leisure is scaled by 25, the average hourly income expected on the middle wage day ($25). This results in the income portion of the function being a larger contributor to the resulting score on high-income days ($30 & $35) and a smaller contributor on low-income days ($15 &

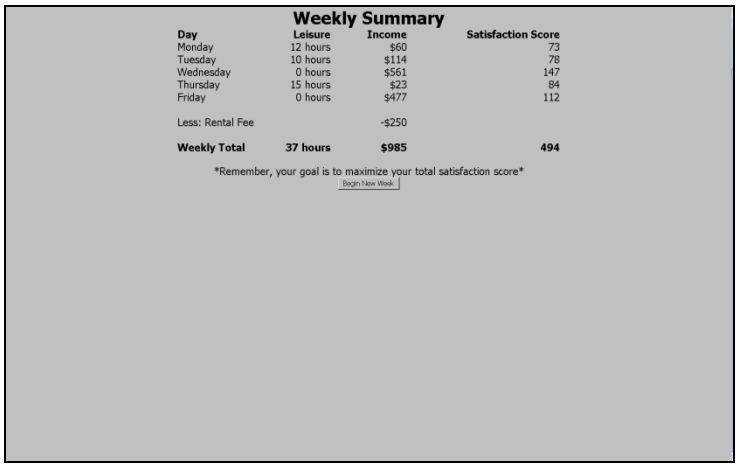
$20) compared to what an hour of leisure would yield. Finally, the subtraction of 115 acts as a scaling value set such that a typical day yielded a satisfaction score of 106 and ranged from roughly 10 to 150. Table 1 illustrates how daily satisfaction scores vary with wage level and number of hours worked. It can be seen that working long hours yielded a greater satisfaction score than working short hours on a good (high-wage) day, and that working short hours yielded a greater satisfaction score than working long hours on a bad (low-wage) day.

Table 1. Examples of Daily Satisfaction Scores

	Hours Worked		
Type of Day (wage)	4 hours	12 hours	Optimal Hours (hours)
Good Day ($35)	114	146	150 (15)
Average Day ($25)	94	94	96 (8)
Bad Day ($15)	73	39	80 (1)

Participants were randomly assigned to one of the two feedback conditions. Those assigned to the daily (i.e., narrow) feedback condition viewed the daily satisfaction score at the end of every day, as well as a summary of daily satisfaction scores, income and leisure hours after every five days (one simulated work "week"), which detailed those same five days. In contrast, participants in the weekly (i.e., broad) feedback condition, only received the summary every five days. Figure 3 shows a screenshot of the weekly summary. Both groups had the objective of accumulating a high total satisfaction score, which was the sum of all the daily satisfaction scores across the session. Participants were informed that they would be paid on the basis of the total satisfaction score they achieved.

Figure 3. Example of weekly feedback summary



Within a given week, wages on any particular day were randomly sampled without replacement from a uniformly-distributed set of five values: $15, $20, $25, $30 and $35. That is, over the course of a simulated week, the participant experienced each of the five possible wage levels for one day each. These daily wage values were used to determine the hourly wages that participant received. In order to have some within-day variability, the wage for a given hour was drawn from a uniform distribution with minimum value of the assigned daily wage minus 3 and a maximum value of the assigned daily wage plus 3. Thus, wages have the property of being uncorrelated across days but correlated within a day, a feature that real cab drivers face (as established in Camerer et al., 1997) and that is important to the assumptions of lifecycle models of labour supply.

Results and Discussion

The main dependent variable in the analyses below is the number of hours worked on a given simulated day. Factors influencing the number of hours worked on a particular day were analyzed using hierarchical linear modeling with restricted maximum likelihood estimation. This method of analysis was used due to the nested nature of the data. All analysis was conducted using the SPSS statistical environment.

The individual-level equation for individual *j* in feedback condition *i* contained three predictors, *wage*, *day* and their *interaction*, where *wage* represents the average hourly wage value for the day and *day* represents the point in the experiment (i.e., the trial or day number, from 0 to 99) on which that wage occurred. This first level equation took the following form:

$$Hoursworked_{ij} = \alpha_{0i} + \alpha_{1i} wage_{ij} + \alpha_{2i} day_{ij} + \alpha_{3i} wage_{ij} day_{ij} + \varepsilon_{ij}$$

At the group-level, the feedback manipulation was included, with feedback being coded 0 for daily feedback and 1 for weekly feedback. As a result, the second level of the model took the form:

$$\alpha_{0i} = \gamma_{00} + \gamma_{01} feedback_i + \mu_{0i}$$

$$\alpha_{1i} = \gamma_{10} + \gamma_{11} feedback_i + \mu_{1i}$$

$$\alpha_{2i} = \gamma_{20} + \gamma_{21} feedback_i + \mu_{2i}$$

$$\alpha_{3i} = \gamma_{30} + \gamma_{31} feedback_i + \mu_{3i}$$

Combining and rearranging the individual and group level models yields:

$$\begin{aligned} Hoursworked = {} & \gamma_{00} + \gamma_{01} feedback_i + \gamma_{10} wage_{ij} + \gamma_{11} wage_{ij} feedback_i + \gamma_{20} day_{ij} \\ & + \gamma_{21} day_{ij} feedback_i + \gamma_{30} wage_{ij} day_{ij} + \gamma_{31} wage_{ij} day_{ij} feedback_i + \mu_{0i} \\ & + \mu_{1i} wage_{ij} + \mu_{2i} day_{ij} + \mu_{3i} wage_{ij} day_{ij} + \varepsilon_{ij} \end{aligned}$$

With the γ's representing fixed effects and the μ's and ε as random effects. The resulting fixed effects are shown in Table 2.

Figure 3 shows the effects of increased wages on the amount of hours spent working for both daily and weekly feedback conditions, near the beginning of the experiment and also near the end once they have had more experience with the task. Although income targeting predicts that a negative relationship should exist between wages and hours worked, it is evident that on average, participants worked significantly more rather than fewer hours on high-wage days compared to low-wage days both early and later in the experiment. In other

words participants' work hours varied with wages as predicted, at least directionally, by intertemporal substitution both at the beginning of the experiment, when they lacked experience with the task and also near the end of the experiment.

Table 2. Experiment 1 Fixed Effect Coefficients

Coefficients	Value	Std. Error	DF	t-value	p-value
Intercept	7.772	1.064	57.895	7.302	.000
Group	-1.418	1.506	57.895	-.942	.350
Wage	.116	.044	57.895	2.668	.000
Day	-.094	.018	57.895	-5.238	.010
Group × Wage	.057	.062	57.895	.933	.355
Group × Day	.038	.025	57.895	1.498	.140
Wage ×Day	.003	.000	57.895	5.841	.000
Group × Wage × Day	-.001	.001	57.895	-1.369	.176

These findings of intertemporal substitution diverge from the behaviour observed by Camerer et al. (1997) in two respects. First, on average participants did not perform income targeting, and instead worked more on high wage days than low wage days. Secondly, this behaviour did not come about after first performing a poor strategy, such as targeting, as was suggested by Camerer et al., but instead was something that was done by inexperienced participants as well. Experience, however did lead to more work being done on high wage than low wage days compared to early on in the experiment. This significant main effect of day and the significant wage by day interaction shown in Table 2 are consistent with the idea that decisions became more optimal with experience.

The between-group manipulation of feedback resulted in no significant performance differences. This could be because it was very easy for participants in both groups to realize each week always contained a very high wage day and also a very low wage day, making the information gained from feedback not helpful in the decision making process. An alternate

hypothesis could be that participants in both groups entered the experiment with a broad (non-myopic) outlook and the manipulations simply had no effect in changing this outlook.

The three-way interaction of feedback, group and day approached significance. As can be seen in Figure 4 participants in the daily feedback condition were more sensitive to low-wage days and overall worked less on them with more experience compared to participants in the weekly feedback condition. However, both groups, regardless of the feedback they were given worked on average similar amounts of time on high wage days. That is, in the group where income targeting would be expected to occur, with narrow feedback being seen as inducing more negative wage elasticities, elasticities were positive and participants in this group were actually somewhat more inclined to work less on low wage days than were those given broad feedback.

The optimal strategies depicted in Figure 3, were calculated by determining what distribution of hours would yield the highest score for a given week, for each of three strategies: maximizing (intertemporal substitution; working more on high-wage days and less on low-wage days), targeting (working less on high-wage days and more on low-wage days) and fixed hours (ignoring the wage and quitting at the same time each day). In this analysis, it is assumed that the daily wage is known from the outset of each day, which is an oversimplification because participants had to estimate the wage by driving. However, this is only really an issue with the optimal intertemporal substitution strategy, which on the lowest wage days suggests working zero hours. Although participants obviously would have to drive more than zero hours to realize it was a bad (low-income) day, they do appear to drive longer on such days than is optimal.

Figure 5 represents the distribution of individual's elasticities. These values were computed by regressing *ln(wages)* on *ln(hours worked)*; the resulting slope coefficient being the elasticity. Here it is evident that targeting was not a common strategy, as almost all slope are greater than 0.

Figure 4. Experiment 1 predicted slopes and optimal strategies.

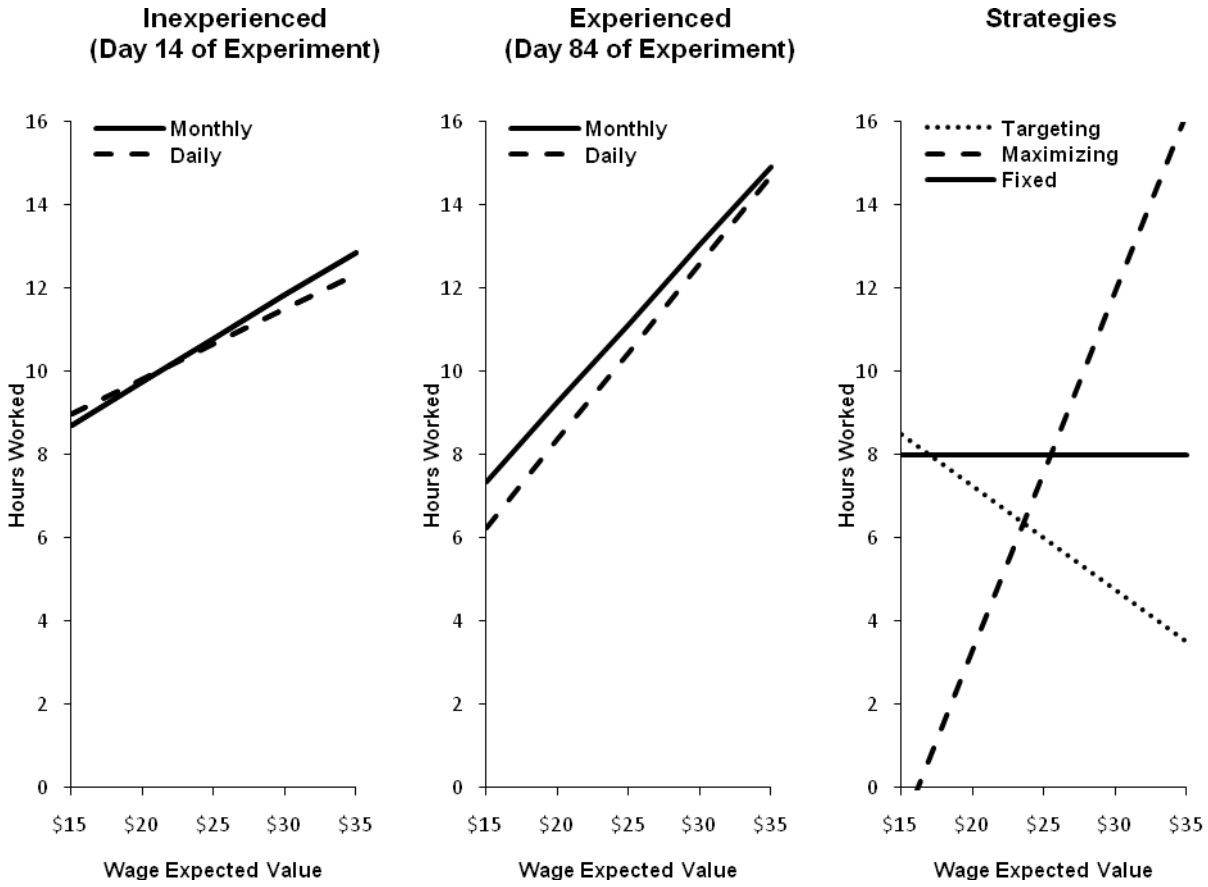
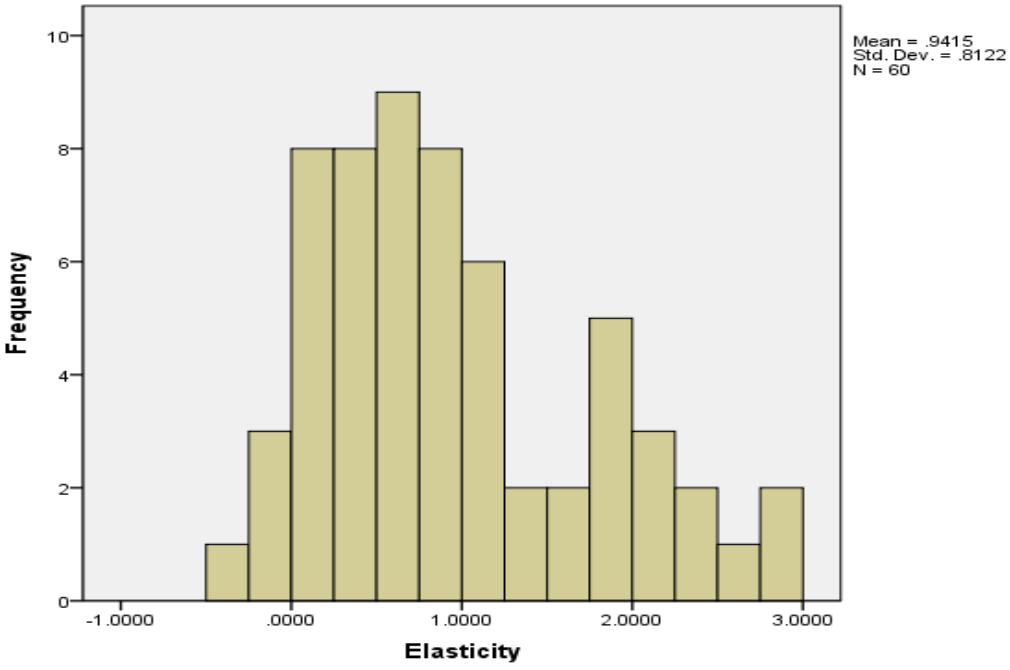


Figure 5. Experiment 1, Distribution of elasticities



Experiment 2

One possible reason income targeting was not a strategy that was common among participants in Experiment 1 was that every week contained good and bad days. It could be that this pattern was easily learned at the start of the experiment and the consistency this provided removed any complexity that would lead to participants performing a simple strategy such as targeting, as the more optimal strategy of intertemporal substitution was fairly easy to execute in the first place. In order to address this possibility, Experiment 2 increased the variability of good and bad days throughout the experiment. That is, while participants still encountered the same overall number of days at each wage level, these wage levels were randomly distributed across the whole experiment, not just within a week. Therefore, compared to Experiment 1, to perform optimally, participants will have to make tradeoffs over larger and less predictable periods. For example, if they encounter a week of low wages, in order to perform intertemporal substitution participants should not work for long periods on any of those days, but instead wait until they encounter days which yield higher wages.

Methods

Participants. Participants were 58 undergraduate students (34 female) recruited from introductory psychology courses, who were told that they could receive between $2-6 for their participation depending on their performance in addition to course credit.

Procedure. With the exception of increased variability of wages between days, the procedure of Experiment 2 is identical to that of Experiment 1.

The base wage value for any day is drawn from a uniform distribution of values: $15, $20, $25, $30 and $35. For both halves of the experiment, consisting of 50 simulated days each, each of the five wage levels were experienced for 10 days each. Within these halves

values are randomly sampled without replacement. In contrast to Experiment 1, then, in Experiment 2 it was possible to experience an unusually good or bad week (i.e., a week with more good than bad days, or vice versa, as defined by average wage level).

Results and Discussion

Similar to Experiment 1, there was no evidence of income targeting being a predominant strategy (Figure 6). There was a significant effect of wages once again, with participants working more on high wage days than low wage days, both earlier on in the experiment as well as near the end of the experiment. However, as is evident in Figure 5, both feedback groups behaved differently later in the experiment, when they had more experience with the task, with regards to how much they worked at the different wage levels. While both groups increased the amount of time they spent working on high wage days over the course of the experiment, those participants who received daily feedback were also more likely to reduce the amount of time worked on low wage days compared to the small decrease in hours worked by those in the weekly feedback condition. Thus, those in the daily feedback condition were once again more sensitive to the daily wage, however this sensitivity was not in the direction we would have expected given the results of Camerer et al.'s study but is consistent with the findings of Experiment 1.

In comparison to Experiment 1 the amount of time worked on low-wage days once participants had become experienced with the task is higher while time spent on high-wage days is somewhat lower. Therefore, by increasing the variability of occurrences of high and low-wage days, participants on average adopted a smaller range of work hours; put differently, work hours appeared less sensitive to wage level in Experiment 2 than in Experiment 1.

It could be argued that participants in Experiment 2 as well as Experiment 1 are not in fact performing intertemporal substitution at all, as they are not being given the ability to transfer income between days, a key feature of intertemporal substitution. As a result the difference in performance between daily and monthly conditions when taking into consideration wages and experience could simply be a result of how participants are learning the task. That is, participants in the daily feedback condition receive feedback at a faster rate, which allowed them to learn faster that working long hours on low wage days is bad and it is better to work more on high wage days. While the weekly feedback condition does eventually receive the same amount of information, its infrequent nature may hinder participant's ability to match the wage quality of a day with the optimal number of hours to work given that wage level to perform well. Therefore it may not be necessary to consider days as being part of a sequence in order

Table 3. Experiment 2 Fixed Effect Coefficients

Coefficients	Value	Std. Error	DF	t-value	p-value
Intercept	7.681	.932	55.768	8.242	.000
Group	-1.325	1.342	55.835	-.988	.328
Wage	.140	.037	56.019	3.832	.000
Day	-.045	.010	56.009	-4.354	.000
Group × Wage	.025	.053	56.052	.475	.637
Group × Day	.053	.015	55.911	3.537	.001
Wage ×Day	.002	.000	56.415	4.328	.000
Group × Wage × Day	-.001	.001	56.394	-2.510	.015

Figure 6. Experiment 2 predicted slopes and optimal strategies

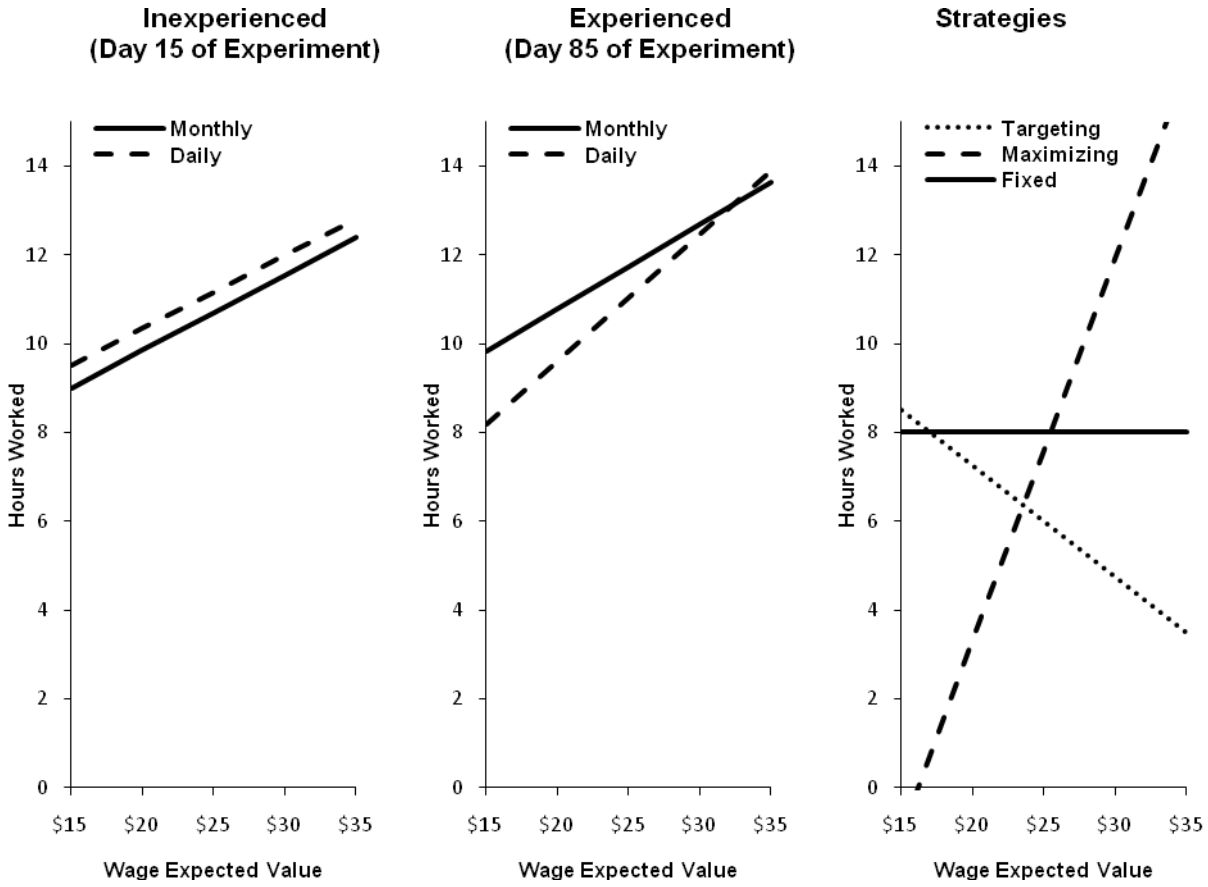
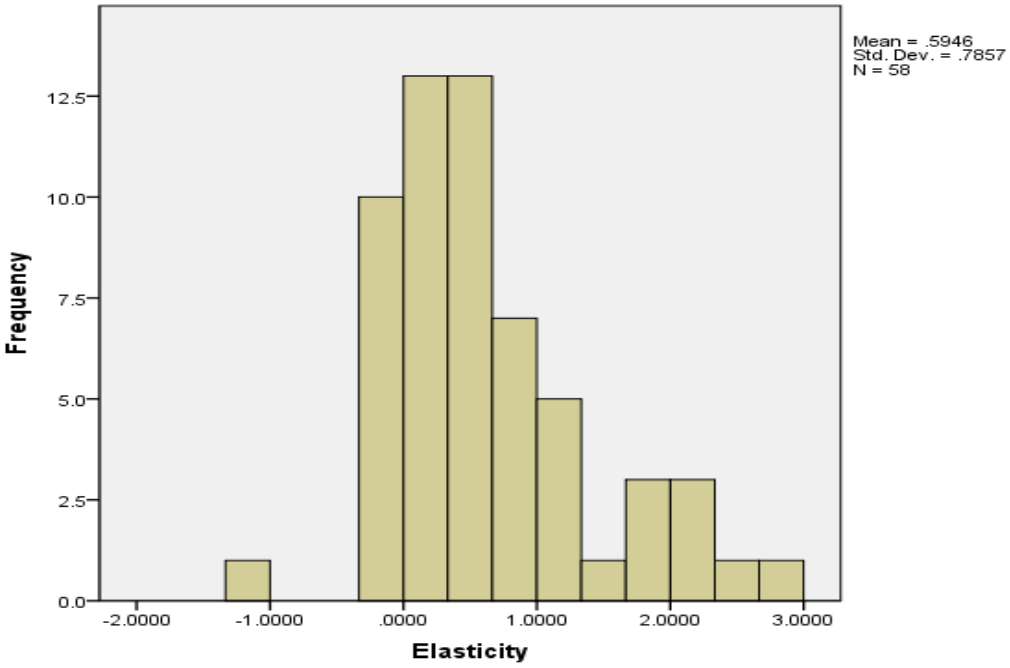


Figure 7. Experiment 2, Distribution of elasticities



Experiment 3

As discussed, Experiments 1 and 2 serve to show that participants are able to distinguish good days from bad and tend to on average work more on high-wage days than low-wage days but are missing a key ingredient of intertemporal substitution. Experiment 3 gives participants the ability to transfer income between days, which under intertemporal substitution would result in the excess income accumulated by working more on high-wage days being spent on low-wage days, in which leisure would be the primary activity for the day. This added choice of when to spend their income may result in making the problem of trying to earn satisfaction more difficult, as now participants have an extra decision to make on top of how much to work on a given day.

Real world cab drivers do not receive a guaranteed income every hour the same way participants did in Experiments 1 and 2. Instead they sometimes must spend periods of time driving searching for customers, which will sometimes lead to them finding one and subsequently receiving a fare. Other times, this searching may not yield any results. Such variability in finding work means that good days are ones in which customers are found more frequently, thus more fares are collected within an hour, and bad days are ones in which few customers are found, meaning fewer fares are collected in an hour.

As a result, Experiment 3 changes the way in which participants pursue work and receive payment. Rather than choosing to work at the hourly level and then receiving a wage for that period, participants will make work/leisure decisions every 15 minutes. When they choose to work, they are in fact searching for customers and have a probabilistic chance of finding one and receiving a fixed fare for that period of time, where the probability is determined by the randomly assigned quality of day. This change was intended to make the

hourly wage less salient and reduce the possibility of demand effects (i.e., subjects seeing wages changing from day to day may feel they are supposed to respond to the change). A by-product of this change is an increase in the computational complexity of the task. By making the hourly wage less salient than the previous experiments, distinguishing "good" days from "bad" becomes more challenging.

Methods

Participants. Participants were 61 undergraduate students (36 female), recruited from introductory psychology courses, who were told that they could receive between $2-6 for their participation depending on their performance in addition to course credit.

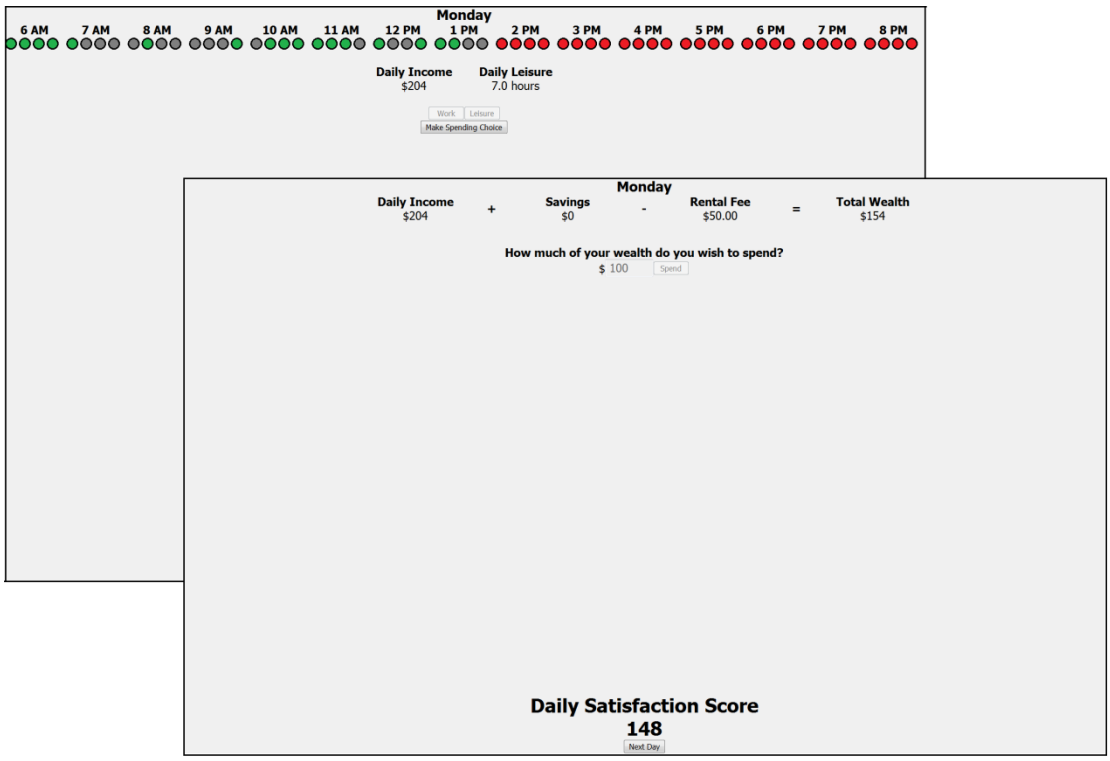
Procedure. Experiment 3 includes the following changes in design and procedure, with all other aspects of the experiment being identical to that of Experiment 2.

Participants completed 80 simulated work days, making four work/leisure choices every hour for 15 hours resulting in 60 work/leisure choices over the course of a day. These choices were displayed as empty circles laid out horizontally across the screen and were filled in with one of three possible colors corresponding to the outcome of a participant's choice: green, grey and red which indicated work was found, no work was found and leisure time respectively.

If the participant chose to work, one of two events would happen: they either failed to find work and received no wage or they found work and received a wage of $12. The probability of finding work for any work/leisure choice within a day was determined by randomly assigning a probability of finding work to that day; the possible probabilities were: 2/8, 3/8, 4/8, 5/8 and 6/8. The occurrences of these probabilities were split evenly across the

first and second half of the experiment, and within these halves they were randomly selected from a uniform distribution without replacement.

Figure 8. Example completed day (daily feedback).

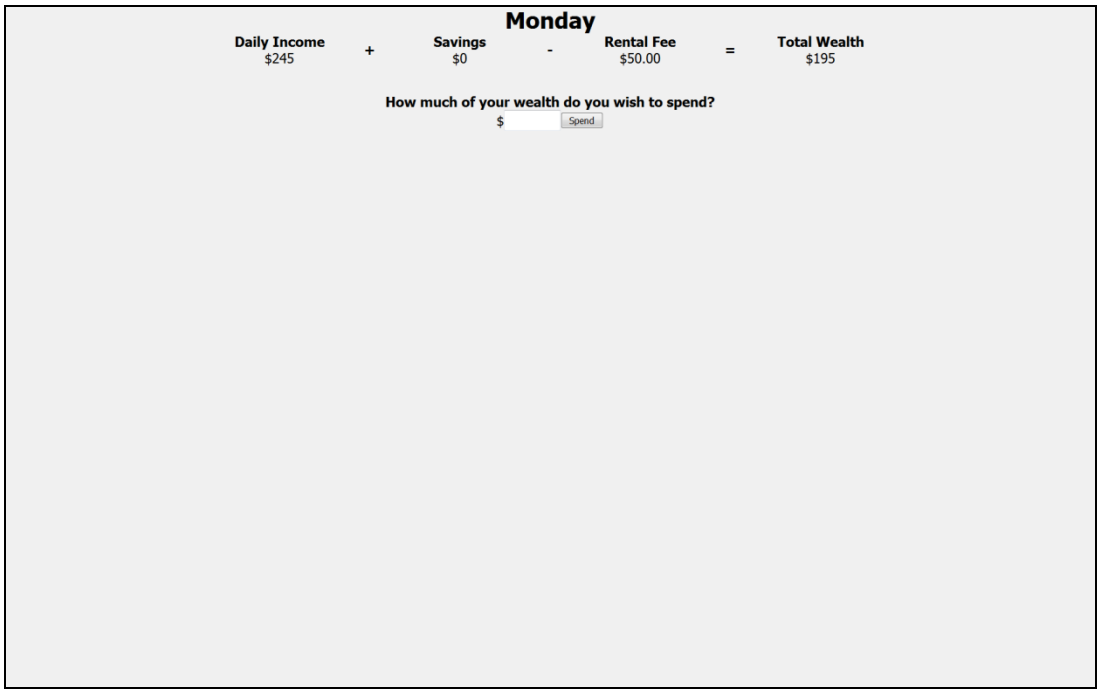


At the end of every day participants were asked, "How much of your wealth do you wish to spend?". Wealth was determined by adding the amount of income earned in the current day with unspent income from previous days (savings), less a $50 daily cab rental fee. Participants could not spend more that their total wealth. If their total wealth was negative as a result of not having enough to pay cab fees, this debt would have to be paid back before they could spend. The amount participants chose to spend was then used along with the amount of leisure hours in the current day to calculate the daily satisfaction score which now took on the following form:

$$daily\ satisfaction\ score = (24Leisure)^{.88} + Consumption^{.88}$$

This daily satisfaction score function contains all the features as the function discussed in Experiment 1, and aside from using consumption instead of income it is only different in how it scales leisure and the final score. These changes were implemented due to the new method of determining wages, which resulted in the average wage on a middle wage day being $24 and overall income throughout an entire day being much lower and thus a naturally lower score is produced without the need for scaling.

Figure 9. Example of spending/consumption screen



Participants assigned to the narrow feedback condition saw their satisfaction score daily, after making their spending choice for the day. After every five days they received a screen that indicated what week they had just finished and a reminder that they would be paid based on their total satisfaction score. They did not receive a summary of their daily activity or scores for the week. In contrast, those participants assigned to the broad feedback condition received no feedback after their spending decision at the end of each day, but instead received it as part of a five day summary presented at the end of each week. As in the narrow feedback

condition, this summary sheet also included an indicator of the current week that was finished and a reminder about pay being linked to total satisfaction score.

Results and Discussion

As in the previous experiments, participants on average worked more on high wage days than on low wage days. However, this relationship's change over time, a reflection of participants' experience with the task, was different. Unlike the previous experiments, experienced participants, regardless of group, did not work more on high wage days compared to early on in the experiment. Participants did however show a reduction in the amount of time spent working on low wage days. This reduction in hours worked on low wage days was much smaller than in previous experiments and was mainly carried out by those in the daily feedback condition. There was, however, no significant difference between feedback conditions.

Despite introducing experimental changes to make the present experiment more similar to the situation faced by New York City cab drivers, such as spending choices and wages more dependent on finding work, income targeting did not become a strategy used by the average participant (Figure 10). That is, despite increasing the computational complexity that participants faced in attempting to earn satisfaction points, participants did not resort to adopting a simple strategy that was suboptimal and instead performed a form of intertemporal substitution. Even inexperienced participants elected to work more on high wage days than low wage days, contrary to the prediction from Camerer et al. that income targeting would be most prevalent in the absence of experience. However, despite the presence of intertemporal substitution participants in both groups did not strongly improve how they set their work hours over the course of the experiment. Rather than resulting in a different strategy being adopted,

increasing the computational complexity of the task resulted in participants not performing intertemporal substitution to its fullest possibility even after becoming experienced with the task. Thus, it appears that participants were simply slower or unable to learn effectively that working even more on high wage days and even less on low wage days would result in them receiving a larger satisfaction score.

Table 4. Experiment 3 Fixed Effect Coefficients

Coefficients	Value	Std. Error	DF	t-value	p-value
Intercept	7.490	.649	59.323	11.536	.000
Group	-.338	.912	59.612	-.371	.712
Wage	.103	.023	59.806	4.472	.000
Day	-.041	.012	58.198	-3.454	.001
Group × Wage	.013	.032	59.972	.397	.693
Group × Day	.024	.016	58.252	1.467	.148
Wage ×Day	.001	.000	58.748	3.526	.001
Group × Wage × Day	-.001	.000	58.703	-1.182	.242

Figure 10. Experiment 3 predicted slopes and optimal strategies.

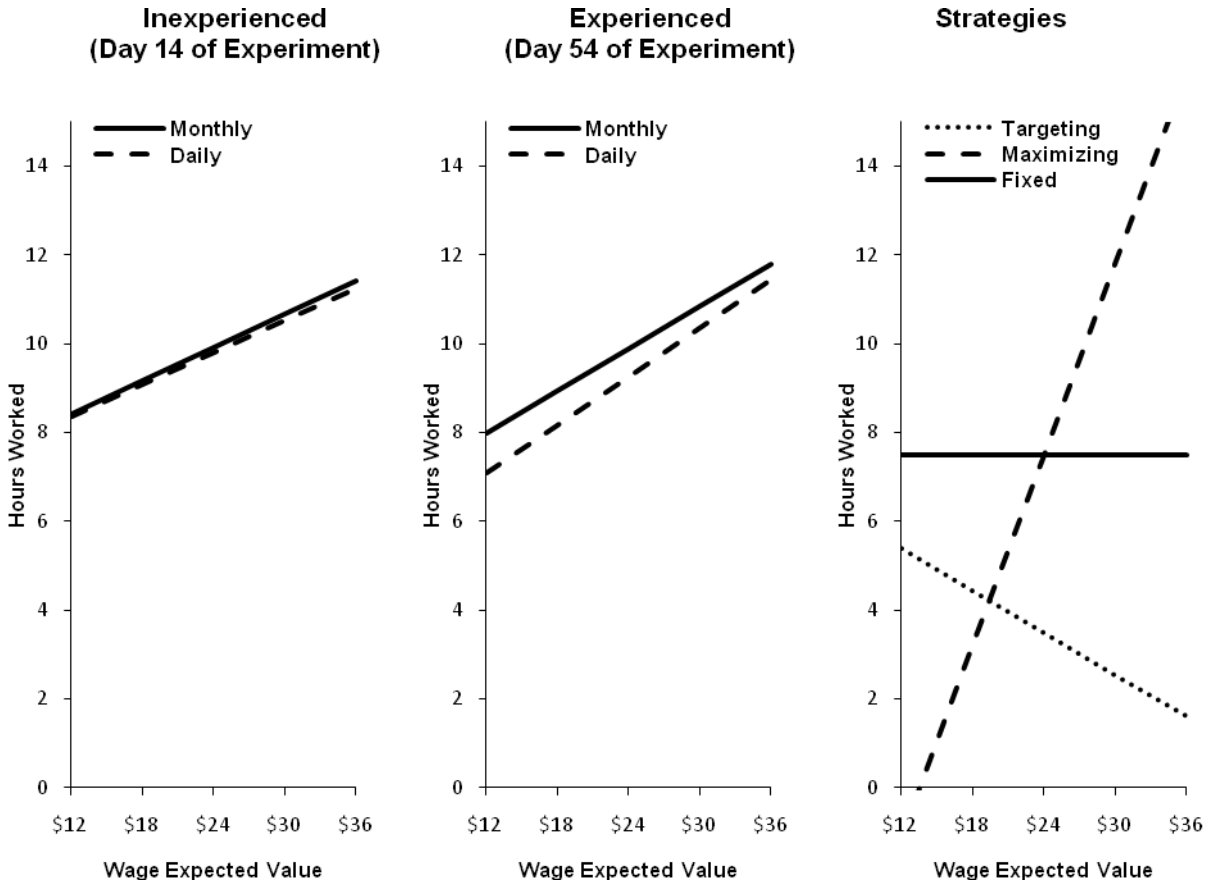
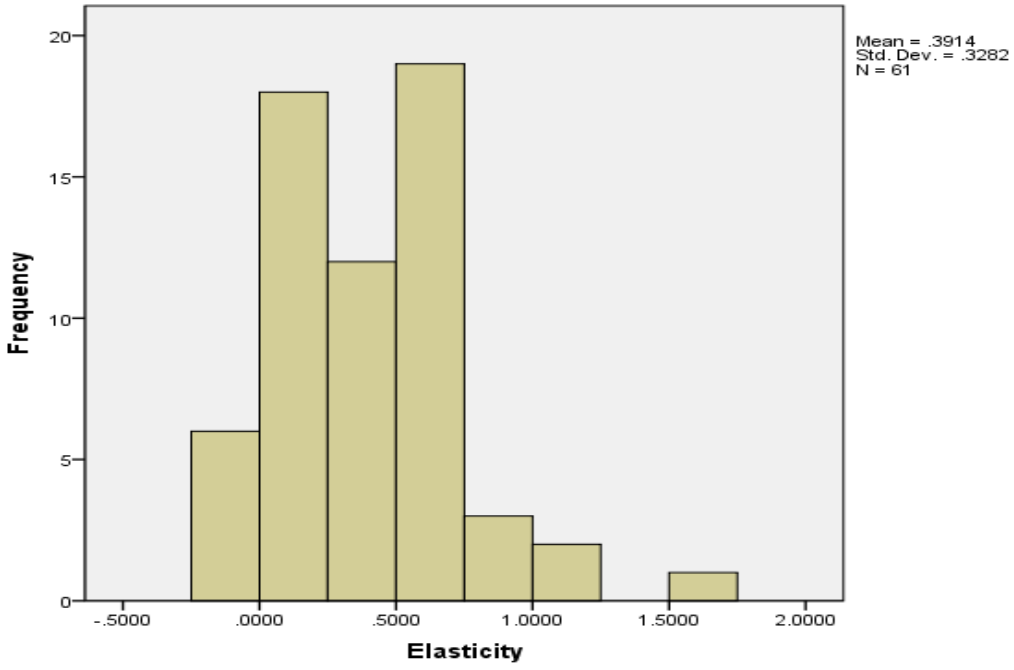


Figure 11. Experiment 3, Distribution of elasticities



Experiment 4

Despite increasing the computational complexity of the strategy required to earn a high satisfaction score in Experiment 3, income targeting was still not used as an alternative, heuristic strategy. While this finding is consistent with what occurred in Experiment 1 and 2 it is important to note that the added complexity did have the effect of impairing the performance and learning of intertemporal substitution. That is, in contrast to Experiments 1 and 2, participants, once experienced with the task, did not work more on high wage days and only showed a small decrease in the amount of time spent working on low wage days compared to when they were inexperienced with the task. As a result we wanted to reduce some of this added complexity while still allowing for the carrying of income across days and a day's wage quality being determined by the frequency in which fares are found.

Experiment 4 accomplishes this by eliminating the spending choice and also reducing the range of different quality days to only two types. Rather than making a spending choice participants are informed that they have a set amount of expenses that are constant throughout the experiment. Paying these expenses becomes necessary to actually receive any satisfaction points earned by participants. These expenses are paid for by accumulating income across two types of days, a low-wage one and high-wage one, which differ in the probability of finding a cab fare when work is taken. Both these changes have the goal of reducing the complexity of the task without removing the new features introduced in Experiment 3. By removing the explicit spending choice participants no longer have a secondary decision to make on top of how long to work on a given day. This change also has the added benefit of allowing more trials to be included in the experiment, thereby giving participants a greater opportunity to

learn the task. Similarly, reducing the number of quality days from five to two should help participants to recognize days as "good" or "bad" more readily.

In addition to these changes the between subjects manipulation of feedback was adjusted. Rather than varying when satisfaction scores are shown, which has been mostly ineffective at this point, possibly due to participants coming into the experiment with a broad perspective that is not changed by the manipulation, the focus will now be placed on expenses. Satisfaction scores will be presented at the end of every day for both groups, however groups will differ in how their expenses and progress towards meeting their expenses is framed. That is, one group will be presented with expenses at a daily level (i.e., narrow feedback) and another will be presented with expenses framed at a monthly level (i.e., broad feedback). While the actual amount of expenses will be the same for both groups, this change is less subtle than the previous between group manipulation. Similar to the original hypothesis stated at the outset of these experiments, it is expected that those participants who have a daily scope for expenses will be more likely to consider each day one at a time, and simply work until their expenses are paid and quit shortly thereafter.

Methods

Participants. Participants were 60 undergraduate students (31female) recruited from introductory psychology courses, who were told that they could receive between $2-6 for their participation depending on their performance in addition to course credit.

Procedure. The experiment consisted of 180 days, each consisting of four work/leisure choices every hour for 15 hours, resulting in 60 work/leisure choices over the course of a day. The 180 days were divided up into six 30 day blocks called months. Within each month participants were tasked with trying to maintain as high a satisfaction score as

they could achieve while ensuring they met their expenses for the month. Participants were still paid at the end of the experiment based on their total satisfaction score. However, if expense payments were not met during a month, the sum of daily satisfaction scores for that month would not be added to the total satisfaction score. Daily satisfaction scores were calculated using the following equation based only on the amount of leisure taken during the day.

$$daily\ satisfaction\ score = 100 - 100\left[\frac{(15 - LeisureHours)^2}{225}\right]$$

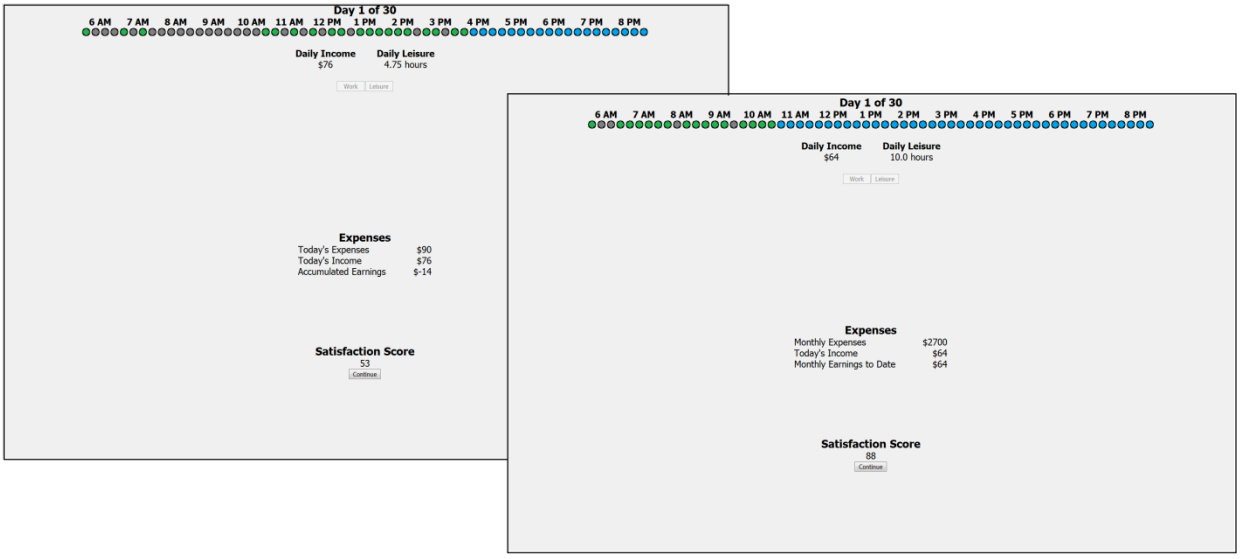
Since the amount of income consumed for expenses is constant, it does not need to be included in the daily satisfaction score function as in previous experiments. This function results in the creation of a range of possible scores from 0 to 100, and is subject to diminishing returns as the amount of leisure hours increases.

In contrast to previous experiments, participants only faced the possibility of two different wage levels, such that there were only bad days and good days. These days were defined by the probability of finding a fare upon selecting to work, on a bad day there was a 40% chance of finding a fare, whereas on a good day the likelihood of finding a fare was 70%. If a fare was found participants received $4 for that period and did not receive anything if no fare was found.

Participants were randomly assigned to one of two groups which differed in how expenses were presented to them (Figure 12). One group was shown expenses as a daily value of $90 and the other group viewed expenses at the monthly value of $2700. At the end of each day, participants in both groups were shown their income for the day as well as their satisfaction score for the day. In the daily condition, that day's expense of $90 was also presented, along with the difference between income and expenses, which could be negative

or positive. This negative or positive balance carried over days within a month, but not across months. Thus in the daily condition, the participant had to have a zero or positive balance at the end of the month to have that month's satisfaction score added to their total for the session. In the monthly condition, the day's income was presented along with the total monthly earnings to date, which was to be compared to the total monthly expenses of $2700.

Figure 12. Example daily (left) and monthly (right) feedback.



At the end of every month (30 days) participants were shown their satisfaction score for the entire month and were informed whether they had met or failed to meet their expenses for the month.

After completing the simulated driving task, participants were then asked to complete two questionnaires. The first questionnaire required participants to accept or reject six lotteries which consisted of the flipping of a fair coin (see Figure 13). Each lottery varied the potential loss ($2-$7) while keeping the potential gain constant at $6. Participants rejecting lotteries with potential losses of less than $6 are considered to be acting in accordance with loss aversion (Gachter et al., 2007).

The second questionnaire asked participants to place themselves on 11 five item likert scales and was used to measure regulatory focus (Higgins et al., 2001). For example, one of the items asks "How often have you accomplished things that got you 'psyched' to work even harder?" These scales are meant to access the motivations behind goal pursuit, either in terms of promotion (seeking success) or prevention (avoiding failure). However, this measure failed to reveal anything of interest and will not be discussed in the results.

Figure 13. Lottery choice task used to measure loss aversion.

Lottery Choice Task

Below are the possible outcomes of a series of lotteries involving the flip of a fair coin. For each lottery please decide whether you would accept it (that is, play it) or reject it (and receive or lose nothing).

Lottery	Accept	Reject
#1. If the coin turns up heads, then you lose $2; if the coin turns up tails, you win $6.	O	O
#2. If the coin turns up heads, then you lose $3; if the coin turns up tails, you win $6.	O	O
#3. If the coin turns up heads, then you lose $4; if the coin turns up tails, you win $6.	O	O
#4. If the coin turns up heads, then you lose $5; if the coin turns up tails, you win $6.	O	O
#5. If the coin turns up heads, then you lose $6; if the coin turns up tails, you win $6.	O	O
#6. If the coin turns up heads, then you lose $7; if the coin turns up tails, you win $6.	O	O

Results and Discussion

As is evident in Figure 14, income targeting was not observed among inexperienced or experienced participants. However, as is evident in Table 5, the interaction between wages and feedback manipulation was statistically significant. Participants who received daily feedback were more likely to work a fixed number of hours despite the quality of the day, whereas those who received monthly feedback about expenses did work slightly more on high wage days. While solely working a specific amount of time no matter the wage quality may seem like a strategy that does not consider multiple days, for the strategy to actually work such that expenses get paid, at some point participants needed to consider what number of work hours would do this. That is, below a certain number of work hours this strategy would fail to generate enough income to pay expenses and result in participants not receiving satisfaction points. Alternatively, participants could have simply learned by trial and error that these hours

resulted in the payment of expenses. Unfortunately, in this circumstance it is not possible to determine if this strategy was learned by considering the tradeoffs between days, by thinking about more than one day at a time, or by trial and error.

Table 5. Experiment 4 Fixed Effect Coefficients

Coefficients	Value	Std. Error	DF	t-value	p-value
Intercept	10.388	1.027	58.612	10.115	.000
Group	-2.575	1.453	58.728	-1.772	.082
Wage	.007	.110	58.708	.068	.946
Day	-.007	.007	57.959	-.996	.323
Group × Wage	.344	.155	58.886	2.216	.031
Group × Day	.001	.010	58.296	.116	.908
Wage ×Day	.001	.001	57.902	.773	.442
Group × Wage × Day	-.000	.001	58.393	-.220	.827

Interestingly, despite the attempt to make the task easier to learn and therefore increase the likelihood of participants adopting a better strategy of work hours with experience, there was no change in participants' behaviour over the course of the experiment. Participants on average used the same strategy that they had used early in the experiment later on in the experiment as well. One potential reason for this could be due to the way in which the probability of finding work and thus a wage were changed. In the previous experiment, the likelihood of finding work ranged from .20 to .80, compared to the current study where the probability of finding work was .40 or .70. Participants may simply be having a difficult time distinguishing good from bad days as a result of the average wages being similar. Therefore, since participants fail to see the high wage day as being better than the alternative low wage day they may not work anymore than they would normally.

Correlating individual's responses on the Gachter et al. (2007) lottery task with their elasticity (slope of hours on wage, see Figure 15) revealed no significant relationship existed (r = .112, p = .416). That is, loss aversion did not seem to play a role in participant's decisions, as would be predicted by Camerer et al., in that under their account people who are

more loss averse ought to be more prone to targeting and therefore exhibit negative elasticities.

Figure 14. Experiment 4 predicted slopes and optimal strategies

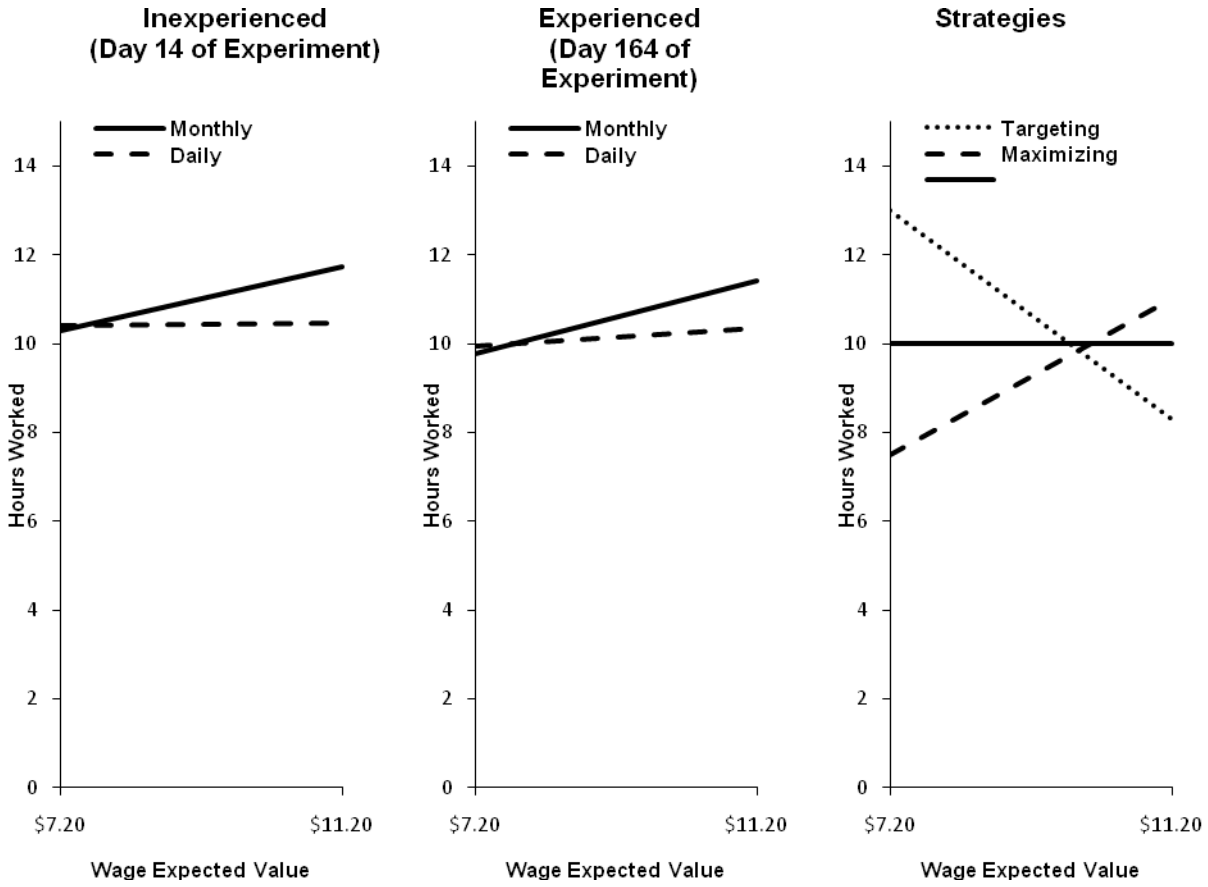
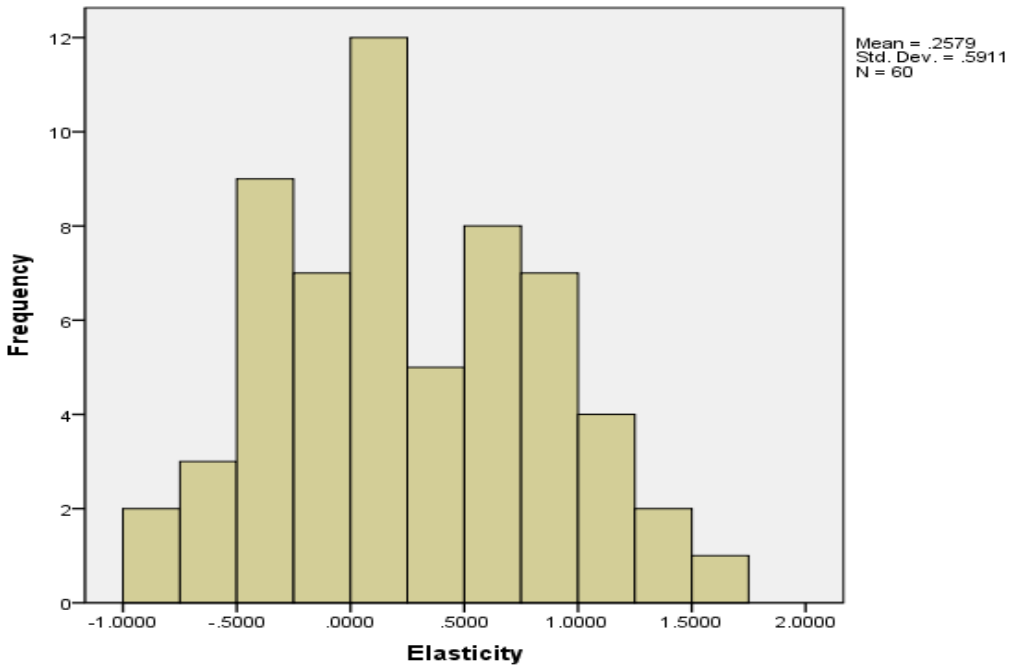


Figure 15. Experiment 2, Distribution of elasticities



General Discussion

Across the four experiments, there was no evidence of income targeting being a predominant strategy. As can be seen in Table 6, despite making various changes to the experimental paradigm, this absence of targeting remained, and instead, intertemporal substitution was most often performed. Based on these findings, it is possible to evaluate the three hypotheses outlined in the introduction, that were largely derived from the field study conducted by Camerer et al..

Table 6. Summary of Experiments and Findings

Exp.	Features	Findings		
		Targeting	Intertemporal Substitution	Learning
1	· 100 days · five wage levels · narrow/broad feedback based on frequency · wages distributed within each week	No	Yes	Yes
2	· 100 days · five wage levels · narrow/broad feedback based on frequency · wages distributed across entire experiment	No	Yes	Yes
3	· 80 days · five probabilistic wage levels · narrow/broad feedback based on frequency · explicit savings choice · wages distributed across entire experiment	No	Yes	Yes
4	· 160 days · two probabilistic wage levels · narrow/broad feedback based on expenses · saving is possible · wages distributed across	No	No	No

entire experiment
- loss aversion questionnaire

H1: Participants will show a negative relationship between hours worked and hourly wage.

The lack of income targeting is perhaps the clearest finding of the present study. With the exception of Experiment 4's daily feedback condition, participants, on average, had a positive relationship between wages and hours worked. Even participants who were inexperienced with the task chose to work more on high-wage days than low wage days. While this positive relationship did decline with each new experiment, it was always significant. The only case in which a positive relationship did not exist was for the already mentioned daily feedback condition of Experiment 4. None the less, targeting still did not occur and instead participants were insensitive to the wage and did not vary their work hours, a possible indication of working fixed hours. A possible reason for the lack of targeting will be discussed shortly. As measured in Experiment 4, loss aversion was not correlated with individual's elasticities. This in itself draws into question the underlying cause of the Camerer et al.'s targeting behaviour.

H2: The relationship between hours worked and wages will be positively moderated by the role of experience.

While on the whole supported, in that participants did primarily improve performance across their time in the experiment, the size of this effect diminished over the course of the experiments, completely disappearing in Experiment 4. It is important to note that this positive moderating effect is somewhat different than the one observed by Camerer et al.. That is, the positive effects of experience observed by Camerer et al. represented the

movement away from a negative relationship between hours worked and wages, thus something resembling more of a strategy change, while the effect observed in the present experiment is more akin to fine tuning a strategy that already existed early on in the experiment. However, as already discussed, the positive correlation observed by Camerer et al. could simply have been due to a confound as a result of the between subject nature of their measure. In a similar field study of cab drivers conducted in Singapore, Chou (2000) failed to find a positive relationship between experience and elasticity, using the same between subject measure of experience. These conflicting accounts could be evidence of a confounding variable, which was not present in our study.

Of note, however, was how this positive effect diminished over the course of the experiments. When reviewing what it was that changed over the course of the experiments, it is evident that while the task took on more features of the conditions of real world cab drivers, such as allowing for savings and wages dependent on finding work, the positive effect of experience on work hours and wages diminished. The simplest solution to why this is the case is that these added features came at a cost of increasing the complexity of the task, which in turn made more difficult the job of allocating time and income optimally. In other words, it would seem that as the task became more realistic, the role of experience diminished. This seems to suggest that, should these findings be generalized to workers in an actual labour market, who face fluctuating wages and are free to choose their hours, there would be a minimal effect of experience on changing the amount of time they work at different wage levels. However, it is possible that there are factors that exist in the actual environment that can contribute to making workers' choices less complex. For example, take the issue of identifying whether the work day can be classified as "good" or "bad". Making this

identification is fundamental to performing intertemporal substitution. In the present study, the only indicator of the quality of a day was the frequency with which fares were found, and in Experiment 4 manipulating the probabilities of "good" and "bad" days had no significant effect on work hours, despite a difference of .70 and .40 respectively. Had another cue also been present to better signal the quality of day, it highly likely that participants sensitivity to wages would be greater. Gureckis and Love (2009), found that perceptual cues that identify the underlying state of a dynamic environment (in the present study's case, good or bad days) aid in the adoption of optimal choices as they deal with the problem of perceptual aliasing, helping to differentiate different states. In an experimental setting, this could be done by simply having a unique identifier appear before each day (trial) and participants learn to associate that identifier with a specific underlying state. In a real cab driving labour market, such an identifier could be weather; likewise in other labour markets, such as vendors at a baseball stadium, information about the opposing team and their popularity serve as an identifier (Oettinger, 1999).

H3: Relative to those given narrow feedback, participants given broad feedback will move away from a targeting strategy toward intertemporal substitution more rapidly as a function of experience.

It is not possible to evaluate this hypothesis as targeting was not a predominant strategy, even during the beginning of the experiment, and between groups, thus, no shift could be made to intertemporal substitution. This could be because participants entered the experiment with a broad perspective in mind already, and the narrow feedback condition was unable to change this. A possible contributor to this issue could be the short time horizon of the experiment, about 1 hour, that participants were aware of before beginning the experiment.

It is possible, that knowing this duration of the experiment made it easier to view the task in a broad perspective. In contrast, actual cab drivers lack this limit on their experience, they do not know when they will no longer be cab drivers, thus their task is much more abstract in terms of duration.

Instead of framing participants' representations of the task into narrow and broad outlooks, it is likely that the differences that are being seen with respect to the feedback manipulation, are simply due to the different intervals of feedback influencing how fast participants learn what allocations of time and income are best. This offers a simple explanation as to why the daily condition in Experiment's 1-3 appear to work more on high wage days and less on low wage days with more experience; with more feedback they have more experiences to draw from in making their decisions. Although, the weekly condition did eventually receive information about all the days, the grouped nature of the data would make it harder to associate the outcome of a given day with the salient experience of what they did.

Experiment 4 deviates from the other experiments in this regard, possibly due to the creation of a goal that is more salient than that of getting a high satisfaction score. By giving expense budgets (daily or monthly), participants may have prioritized ensuring these budgets were met without regards to their satisfaction scores. This would have resulted in each feedback condition performing a different strategy with respect to the budget. Those in the monthly condition, could be simply trying to hit that target as soon as possible and then take the rest of the time in the month off, as additional work hours would not be useful. In contrast, participants in the daily condition may have a better idea of where they are with respect to their expenses and do not attempt to build up a reserve of funds.

A lack of income targeting and next steps

While targeting was proposed to be driven by myopic loss aversion, something that was not apparent in this study, the reason that this heuristic was used by the cab drivers, according to Camerer et al., was to solve problems of computational complexity and self control associated with performing an optimal strategy such as intertemporal substitution. As Camerer et al. describe it, targeting is a simple decision rule that only requires keeping track of income, whereas, intertemporal substitution requires keeping track of forgone leisure and income and the resulting marginal utilities. Targeting was also proposed by Camerer et al. to help mitigate self control problems associated with intertemporal substitution. To successfully carry out this strategy, drivers have to avoid the temptation of spending excess income on high wage days and instead save it for use on low wage days; and, perhaps more critically, they have to avoid the temptation to quit early on high wage days (a form of procrastination).

The present study only placed emphasis on and tested the computational part of this problem. Participant's goal of obtaining a high satisfaction score was simply an optimization problem, and while the success at performing this optimization varied across experiments, at no time did targeting become a strategy used to avoid solving this problem. If anything, learning the underlying dynamics of the task provided the most difficulty for participants. Our our variations of complexity didn't lead to targeting, which suggests that the key difference between the task in the present study and Camerer et al.'s cab drivers, is more associated with self-control.

As a result, computational complexity by itself can be ruled out as the cause of income targeting. Ruling out task complexity leaves only self-control issues being a driving cause of the use of income targeting as a heuristic. Our task did not pose any strong self control

challenges: Participants could not buy anything tangible with their earnings in the simulation, there was no temptation to spend earnings that should be saved; and working was not actually more effortful or aversive than simulated leisure, so there was no temptation to quit early on a given day and promise oneself to work longer hours the next. These elements were deliberately excluded from the experiments, with the purpose of disentangling the root causes of income targeting. Had this study contained an effort component in addition to the present design it would not be possible to attribute the cause of targeting to computational or self control issues.

Having ruled out computational complexity as being the sole contributor to income targeting, future studies need to examine the role of self control in determining the relationship between wages and hours worked. The easiest method of doing so would be to add an effort component to the work task. While there exists field studies looking at the role of effort at work (Oettinger 2001; Locke & Mann 2009, Goette, Huffman & Fehr 2004), these tasks primarily focus on how effort itself varies with respect to changes in wages. Dickinson (1999) further analyzed the effort and hours worked relationship using an experimental work task that spanned multiple days and attributed the negative elasticities to the substitution of on the job leisure with off the job leisure.

Conclusion

This study sought to better understand how people make sequential decisions under conditions of uncertainty, by constructing a lab based equivalent of Camerer et al.'s New York City cab driver field study. While cab drivers appeared to be considering one day at a time and taking a narrow perspective with respect to their labour/leisure decisions, the present study revealed that participants, regardless of experience level, made decisions more congruent with

having a broad outlook. Rather than treating days as separate isolated events, participants effectively if non-optimally opted for a more integrative choice strategy that incorporated beneficial tradeoffs across the decision sequence.

References

Benartzi, S., & Thaler, R. (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics, 110*(1), 73-92.

Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). *Labor supply of New York City cabdrivers: One day at a time*

Chou, Y.K. (2000). Testing alternative models of labor supply. Evidence from taxi-drivers in Singapore. *Department of Economics – Working Papers Series 768*, The University of Melbourne.

Dickinson, D. L. (1999). An experimental examination of labor supply and work intensities. *Journal of Labor Economics, 17*(4), 638-670.

Eriksen, K., & Kvaloy, O. (2010). Myopic investment management. *Review of Finance, 14*(3), 521-542.

Farber, H. (2005). Is Tomorrow Another Day? The Labour Supply of New York City Cabdrivers. *Journal of Political Economy, 113*(1), 46-82.

Gachter, S., Johnson, E., & Herrmann, A. (2007). Individual-level loss aversion in riskless and risky choices. Available at http://www.nottingham.ac.uk/cedex/documents/papers/2007-02.pdf.

Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics, 112*(2), 631-645.

Goette, L., Huffman, D., & Fehr, E. (2004). Loss aversion and labour supply. *Journal of the European Economic Association, 2*(2; 3), 216-228.

Gureckis, T., & Love, B. (2009). Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition, 113*(3), 293-313.

Haigh, M., & List, J. (2005). Do professional traders exhibit myopic loss aversion? An experimental analysis. *Journal of Finance, 60*(1), 523-534.

Heath, C., & Larrick, P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology, 38*(1), 79-109.

Higgins, E., Friedman, R.S., Harlow, R.E., Idson, L.C., Ayduk, O.N., & Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology, 31*(1), 3-23.

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

Kliger, D., & Levit, B. (2009). Evaluation periods and asset prices: Myopic loss aversion at the financial marketplace. *Journal of Economic Behavior & Organization, 71*(2), 361-371.

Larrick, R., Heath, C., & Wu, G. (2009) Goal-induced risk taking in negotiation and decision making. *Scoial Cognition, 27*(3), 342-364.

Locke, P., & Mann, S. (2009). Daily income target effects: Evidence from a large sample of professional commodities traders. *Journal of Financial Markets, 12*(4), 814-831.

Mehra, R., & Prescott, E. C. (1988). The equity risk premium: A solution? *Journal of Monetary Economics, 22*(1), 133-136.

Oettinger, G. S. (1999). An empirical analysis of the daily labor supply of stadium vendors. *Journal of Political Economy, 107*(2), 360-392.

Read, D., & Loewenstein, G. (1995). Diversification bias - Explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology-Applied, 1*(1), 34-49.

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