Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Laurens Verhoef

The Effects of Labeling on Consumption Behavior

MSc Thesis 2010-041

MASTER THESIS ECONOMICS

The effects of labeling on consumption behavior

By
Laurens Verhoef (527223)
MSc Economics

Supervisor:
Prof Dr. H.M. Prast

Word count: 14558

August 2010

Table of contents

1. Introduction

Individuals do not always act in line with the rational choice model. Decisions they make can be biased by the context in which the decision is made. Changing the way a choice is presented or the environment in which a choice is made can alter people's decisions. This is called the framing effect. One form of framing is labeling: a label attached to a product or income may influence consumer behavior. There are examples of labeling all around us. Some labels are deliberate: just think of the logos indicating the healthy choices in a supermarket or the mandatory risk indicators of mutual funds. These labels are designed to help us make better choices. More often, labels are not deliberate, and those attaching them may not realize that they influence choice: think of the label "child benefits" on income. This label has the potential to influence the way the income is spent (Kooreman, 2000). Personally, I do not notice that my behavior is influenced by those labels. This made me wonder how effective labeling is as a mechanism to influence choice. A paper by Abeler & Marklein (2008) seems to give an answer to this question. They have shown that when a label is attached to an income source to encourage consumers to spend this income on one particular good, these consumers tend to spend more on that particular good than the control group, that received this income without a label on it. This observation is called the labeling effect. This labeling effect persists even if this extra income should not change choices of rational individuals. Therefore labeling can be used as a mechanism to influence consumer behavior. However, quite often labels are not even meant to influence consumer behavior, while they do. Therefore it is important for all parties that use labeling to be aware of all potential consequences and make sure they use the mechanism properly.

Abeler & Marklein did two experiments to prove the presence of a labeling-effect: one field experiment and one lab experiment. Their main conclusions are that a) a labeling effect is present in both experiments and b) subjects with high mathematical ability are less sensitive to this labeling effect. However, while reading their paper, I found two things that made me question their conclusions. Their research methods could have biased the results in my opinion, overestimating the labeling effect. Their field experiment was conducted in a wine restaurant, which attracts a different crowd than a normal restaurant. Moreover, for their lab experiment they did not include economists in their sample (p. 10). Considering the fact that Abeler & Marklein also find that subjects with higher mathematical ability are less sensitive to the labeling effect, this procedure could have led to an overestimation of the labeling effect.

It has been shown before that economists sometimes act different from non-economists in economic experiments. Carter & Irons (1991) found that, compared to non-economists, economists tend to act more like a "homo economicus": a rational and utility-maximizing individual, who only cares about self-interestTherefore I think it would be interesting to investigate whether economists act significantly different from non-economists in an experiment like the one Abeler & Marklein did and whether their

conclusions regarding the labeling effect are robust to the inclusion of economists in the sample. To test this hypothesis, I have conducted an experiment similar to the lab-experiment by Abeler & Marklein. In addition, I added an extra treatment to my experiment to test whether it makes a difference whether the consequences of a choice are framed as gains or losses. This is called a framing effect and this effect was first shown by Tversky & Kahneman (1981).

The main goal of this thesis is not to question the presence of a labeling effect among non-economists, but to test whether the conclusion that a labeling effect is present is robust to including economists in the sample and to investigate whether economists are less sensitive to the labeling effect than non-economists. The main question that is going to be answered in this thesis will be:

Are individuals sensitive to labeling, and if so, are economists less sensitive to the labeling effect?

Note that in this thesis the term "economist" will refer to a student with economics as his or her major, since only students and recent graduates participated in the experiment that was conducted to test this hypothesis. The reason why economists might have less of a labeling bias is that they are much more calculating than non-economists and used to maximizing utility. It would be easier for them to analyze an optimization problem and find the optimal consumption bundle. Existing research has already proven that individuals can be sensitive to labeling and framing, but this research question is still very relevant because it will test whether a certain group responds differently to government manipulation. If this is the case, then this knowledge can be useful for a government, especially when it comes to manipulating behavior of companies or banks, which are usually led by economists.

In the experiment, it will be tested whether economists act differently from non-economists and whether this is due to their economic education or due to other factors. This can be tested by looking at the difference between economists near the end of their studies and economists that have just started their studies. Furthermore it will be tested whether a labeling effect is present. In that case, it would be interesting to see if economists are less sensitive to this labeling effect. Finally, it will be tested whether it makes a difference for a subject's rationality if the consequences of a choice are gains or losses.

The structure of this thesis will be the following:
- Section two will be the theoretical background. In this section, I will look into theory of behavioral economics and explain when and why subjects deviate from the rational choice model. Also labeling, framing and the effects these mechanisms can have on consumption behavior will be explained in detail.
- Section three will provide the explanation of the experiment, its goals and the expectations about the outcome of the experiment.

- Section four will present the results of the experiment and will also provide a first explanation of these results.
- In section five the implications of the results of the experiment will be discussed. I will first compare the results with existing literature and after that I will discuss the policy implications of the results.
- Section six will provide an overall conclusion.

2. Theoretical background

It is impossible to find one model that exactly describes how consumers make their choices. The most widely used model to predict consumer behavior is the rational choice theory. This theory assumes that consumers always seek to maximize their (long-term) well-being and make their decisions accordingly. This may not always describe actual behavior. Samuelson (1937) assumes that individuals maximize well-being in a time-consistent way: if an individual decides today how to allocate consumption over a certain number of periods, then he/she wants to stick to that plan in the next period. However, humans are not always time-consistent; there exists a present bias. Samuelson was aware of this and pointed out that his assumption may not describe actual behavior. The present bias can be included in a model by using hyperbolic discounting (Loewenstein & Prelec, 1992; O'Donoghue & Rabin, 1999). This thesis will however not focus on the present bias and self-control problems in intertemporal choice, but rather on other biases that can cause individuals to deviate from the path that maximizes their welfare, even when they do not face an intertemporal problem. It will be shown that even in a simple setting where utility should be maximized, people are not always as calculating as economic theory seems to assume.

This section will present the theoretical view on individual decision making and some of the biases that can influence choice behavior. As it is the main subject of this thesis, the focus will be on the effect of labeling on consumption. In order to evaluate to what degree labels cause people to deviate from what would be optimal for them, it is first necessary to define what is considered to be optimal. The most applied theory when analyzing decision making and choice behavior is the rational choice theory. However, this theory needs a lot of assumptions that may not always be satisfied. This problem will be discussed in section 2.1.

2.1 Rational choice theory

In traditional economic theory an action is considered to be rational if it balances (expected) benefits and costs in such a way that it maximizes personal advantage (Friedman, 1953). The assumption that all individuals maximize their benefits (or utility) is the main assumption in the rational choice theory. In order to be able to maximize, one needs to be aware of all available options and their consequences. These consequences can either be a certain outcome or a set of uncertain outcomes with a corresponding probability.[1] However, it may not always be the case that individuals are fully aware of all options and/or consequences. The concept of bounded rationality (Simon, 1957) provides an explanation for this. Individuals may not always have access

[1] Note that the potential outcomes do not always have to have a weight equal to the corresponding probability. The prospect theory by Kahneman & Tversky (1979) uses weight as a function of the probability to calculate expected utility. This is considered to be a more realistic alternative to traditional expected utility theory.

to all available information, they may not always know about all available options or they lack the ability to process all available information in a proper way. We can only predict that an individual will choose the best option out of all options that are known to him, using all accessable information (Jacoby, 2000). Therefore the assumption of full rationality is one weakness in rational choice theory.

Another problem in the rational choice theory, related to bounded rationality, is that in some cases the search costs of acquiring additional information do not outweigh the benefits. For example when you want to buy a snack at a large train station, you are most likely to buy it in one of the stores available at the train station. You will not check the prices and available items in stores outside of the station before you buy your snack, simply because this is too time consuming. In economic terms: you deliberately choose from a limited set of options, instead of searching for all available options. Since it could be the case that the option that would maximize utility was not available in this set of options, the subject does not always maximize utility. Rather than searching for the *best* option, the subject searches for an *acceptable* option. In economic literature, this is called "satisficing" (see, for example, Krosnick, 1991). This would be a more realistic view on consumer behavior than assuming that consumers will always maximize utility.

The main explanation for this phenomenon can be found in the presence of information costs. A potential solution for a paternalistic government would then be to supply information to consumers, to reduce these information costs. Providing information can be done through marketing campaigns but also through labeling. If consumers are unable to make the optimal decision themselves because of a lack of information, the government may be able to push them towards a better decision by attaching labels to products or income. This would reduce the information problem and lead to a higher level of well-being. However, not all labels are designed only to change decision behavior. More on this subject will follow in sections 2.4 and 2.5.

There are many more assumptions of the rational choice theory that can be questioned.[2] The most relevant ones for this thesis are the maximization assumption, the information assumption and, to be discussed later in this section, the fungibility assumption. The most important conclusion from this section is that the rational choice theory predicts that consumers will always choose the option they like best out of a given set of options, regardless of context. In the next section it will be argued that the way a choice is presented or the environment in which a choice is made matters for the final outcome. In other words: consumer behavior is sensitive to framing.

[2] For more information on this subject, see for example Jacoby (2000).

2.2 Framing

People often do not interpret choice situations in the same way as economists generally assume. Features of the environment that should not have any influence on the final outcome, such as the way the options are presented, can affect choice behavior. Tversky & Kahneman (1981) were among the first to show this framing effect in an experiment. They introduced subjects with a hypothetical situation in which there is an outbreak of a deadly disease, which would kill 600 people if no action is taken. They asked subjects to evaluate two programs that have the potential to cure this disease. Subjects were told that program A will save 200 lives with certainty and program B will save all 600 lives with a probability of 1/3 and save no people at all with a probability of 2/3. In this setting, 78% of all subjects preferred program A. However, if they asked subjects to evaluate program C where 400 people would die with certainty and program D where nobody will die with probability 1/3 and 600 people would die with probability 2/3 in the other program, they found that a majority of the subjects suddenly prefers program D, even though this was the same program as program B. The difference can be explained by the fact that people are generally risk averse when it comes to choices involving gains and risk taking when it comes to choices involving losses. Obviously this is not a stable preference and a violation of the rational choice theory, which would predict that a subject would always choose the most preferred option regardless of context. This is an example that shows that the choice made by an individual (revealed preference) is not always equal to the choice that he actually considers to be the most desirable (normative preference). Beshears, Choi, Laibson & Madrian (2008) identify five factors that increase the likelihood that revealed preferences are not equal to normative preferences: passive choice, complexity, limited personal experience, third-party marketing and intertemporal choice. In these cases subjects may be uncertain about their competency to make a good decision and are therefore more likely to be biased by framing.

Also impulsivity could be a sixth factor that can change consumption behavior. Consider the example of a school cafetaria, as explained in Thaler, Sunstein & Balz (2010). A school director knows that final choice of his students depends on the order in which the food is displayed. In this case it is not possible to present this choice in a neutral environment, there has to be some kind of order in the way the products are presented to the students. The director then has several options at his disposal. Financially it would be the most appealing to him to arrange the food in such a way that his profits (or the sales of suppliers that pay the highest bribe) are maximized. He could also arrange the food in a way that makes his students the best off, which obviously would be the socially desirable thing. A third option would be to display the products in a random order. The director in this example has the role of a choice architect. This choice architect has the responsibility to control the environment in which a choice is made. A government often takes the role of choice architect too. They also influence the environment a choice is made, by attaching labels to income or simply by displaying information in a certain way. This will be discussed in more detail in section 5.

2.3 Labeling

In several studies it has been observed that putting a label on a certain option or giving a discount that has to be spent on a certain product is an effective way to influence consumer behavior (Abeler & Marklein, 2008; Epley, Mak & Idson, 2006; Hoynes & Schanzenbach, 2009; Kooreman, 2000). This could have two effects on well-being: either a positive or a negative effect. A positive effect due to labeling could be observed in case consumers make a suboptimal choice in the situation without labeling and correct this behavior because of labeling effects. This could occur in the health domain, where labels are currently used in The Netherlands to promote a healthy food choice, but it can also be very useful for a government to encourage savings for retirement (Card & Ransom, 2007). In order to use labeling in an optimal way, it is necessary to fully understand what effects labeling could have on choice behavior and which groups are the most effected by this mechanism. If some groups are relatively unaffected, this could mean that it is possible for a government to make use of asymmetric paternalism.

The negative effect of labeling could occur if in the initial situation the consumer chose one option, but chooses a worse option because of the labels that are used. An example of this would be to give discounts on one or more products. Even though a consumer could get lower utility from buying this particular product on sale, he would still be tempted to buy it just because it has the label “discount” on it. A numerical example: you are at the supermarket thinking about buying a drink. In advance you know you restrict yourself to two options: brand A and brand B, where brand A gives you utility of 50 and brand B gives utility of 40. In words: you would prefer brand A over brand B. Now you enter the supermarket and find out you can get a discount if you buy brand B. Because of this discount your utility from buying brand B will raise with 5, so total utility from buying brand B would become 45. This is still lower than the utility you would get from buying brand A, so any rational consumer would buy brand A. However, many consumers change their choice to brand B just because of the fact that they will get a discount when they buy this product. Existing research supports this conjecture. Gupta (1988) develops a model on the impact of sales promotions on consumer decisions and finds that 84% of the increase in sales due to promotion comes from brand switching. The extra demand of existing customers (16%) because of this promotion is relatively small. This study shows that knowledge about labeling can also be used by marketeers. Labels, such as discounts, can influence consumer choice and can actually leave consumers worse off than they were without a discount. Therefore labeling can also be used by firms to increase sales (while harming consumers).

Nontheless, McClure et al. (2004) measure brain responses and argue that brand labels on cola also influence consumer preferences. According to Beshears et al. (2008:8) this does not need to be a problem because it wouldn’t harm consumers. If people feel better when they drink one brand, even though they prefer the other brand in blind tests, the brand label doesn´t leave them worse off. The discussion if and how much third party marketing harms consumers is not very relevant for this thesis though. The

most important thing to take from this section is that labeling is not only for governments that want to help citizens make better choices, but that it can be used by marketeers as well. In that case, the interest from the sender of the label is often not the same as the consumer. This possibility should not be neglected.

2.4 Labeling and the fungibility assumption: a survey of existing literature

As shown in the previous section, labeling on products can influence consumer choices. This section will show that labeling on income could matter for the way the income is spent. The rational choice theory implies fungibility of money: the context in which income is received or the source of the income should not matter for the way it is spent. Soon we will see how labeling can cause a violation of this assumption.

Examples of labeling can be found everywhere around us. These labels (sometimes unnoticed) influence our consumption behavior. Kooreman (2000) found evidence supporting the labeling effect of the child benefit system. In his paper, he uses expenditures on both child and adult clothing as dependent variable, where the amount of child benefits received and the income received from other sources serve as an explanatory variable. The fungibility assumption implies that the source of income should not matter. If this assumption would hold, the coefficients of these two explanatory variables should be the same. For households with one child aged 0-11 however there is a significant difference between these two coefficients. The results imply that if child benefits were to rise by 1 guilder, around 11 cents are used to buy child clothing. If other income would rise by 1 guilder, only 1 cent would be spent on child clothing. This is a serious violation of the fungibility assumption. Some more research on this topic has been done in other countries. For example, Blow, Walker & Zhu (2006) find the exact opposite result for child benefits in the United Kingdom: their conclusion is that parents who receive child benefits spend relatively more on adult goods, such as alcohol. Lyssiotou (2009) adds to this discussion by discussing gender differences in the labeling effect of child benefits.[3] She finds that households where the mother receives the payment spent more on child goods than families with children that did not receive the payment. This seems to be in line with the labeling effect. However, households where the father received the payment were observed to spend relatively more on tobacco than households that did not receive the payment. Both findings support the non-fungibility of child benefits, but only the households where the mother received the benefits show evidence for the labeling effect of child benefits. It seems likely though that this gender-difference in the sensitivity to the labeling effect does not hold in other cases, but this can be tested in the experiment.

[3] Lyssiotou did her research on Cyprus, where before 2003 only families with at least four children received child benefits. Families with fewer children were used as a control group. Since there might be a selection bias present here and the sample size is not that large (55-101 observations in the treatment group over the years), the results might not be that valuable, but the conclusion of the paper is nonetheless quite interesting and worth mentioning here.

In field experiments it may be hard to get good data on expenditures and to isolate the effect of the labeling. There will always be unobserved factors influencing the results. Abeler & Marklein (2008) did both a field experiment and a lab experiment on the effects of labeling on consumption. In their field experiment it was tested whether a label on a discount voucher in a restaurant made any difference. In a wine restaurant where guests always ordered for more than 8 euros on beverages, guests were given a discount voucher after entering the restaurant. There were two kinds of vouchers: the guests either received a voucher of 8 euros for the whole bill (cash treatment) or a voucher of 8 euros that had to be used for beverage consumption (label treatment). Obviously guests were only handed one voucher, so they participated in only one treatment. Considering the fact that guests always spent more than 8 euros on beverages, the fungibility assumption would imply that this label should not have any effect on the consumption decision. The first conclusion by Abeler & Marklein was that guests in both treatments spent more money because of this voucher. This is not a surprising result; other research has shown this effect before (See, for example**,** Milkman & Beshears, 2008). What is special and violates fungibility is that guests in the label treatment spent relatively more on beverages than guests in the cash treatment. This is evidence for the existence of a labeling-effect. This increase in spending did not occur because these guests drank more, but because they order more expensive drinks. Abeler & Marklein conclude from this field experiment that guests did not treat these vouchers as fungible.

To find out more about the forces behind such irrational behavior, they also did a lab experiment on the effects of labeling on consumption. This way it would be possible to gather more information about a subject's background. The setup of this experiment was similar to the field experiment: in both situations a discount voucher (subsidy) was handed out and subjects were randomly assigned to the cash treatment or the label treatment. Consistent with their expectations, their result was that subjects in the label treatment spend relatively more money on the subsidized good. Moreover, subjects with low mathematical ability were more likely to violate the fungibility assumption. This finding is the basis for the hypothesis that economists are less likely to violate the fungibility assumption. If it turns out that economists, who generally have high mathematical ability, act more rational in such an experiment, then it could be that this particular group is less sensitive to labeling.

This raises the question whether it is only personal mathematical ability that affects this sensitivity or if there is also an effect of education. Can scientific thinking be taught? Henk Schmidt, rector magnificus at Erasmus University Rotterdam, asks himself this question in the dies lecture on 6 November 2009. During his lecture, he shows evidence for the statement that education causes a clear increase in the statistical reasoning ability of students in social sciences compared to students of natural sciences over the years. For logical reasoning the opposite is true: students in natural sciences become better at logical thinking, while students in social sciences do not increase this ability during their study. This seems to be evidence for the statement that scientific thinking

can be taught. This hypothesis will be tested in the experiment and the results will be presented and discussed in sections 4 and 5.

2.5 Mental accounting

One reason why individuals often violate the fungibility assumption can be found in mental accounting. Mental accounting is defined by Thaler (1999, p. 183) as "the set of cognitive operations to organize, evaluate and keep track of financial activities". Mental accounting is useful to make complicated decisions in finance easier. For example, consumers may have different mental accounts for spending on different products (or product categories), each with their own implicit or explicit budget. Then, if they get a subsidy on one product but not on the other, this only increases the budget for (and spending on) that one product. These individuals would only consume more of the subsidized product and not change their spending on other goods. Obviously not all individuals who engage in mental accounting are this strict in using their mental accounts. The general assumption is that low-income households tend to have stricter and shorter horizon budgets and therefore are more likely to engage in mental accounting. (Thaler, 1999) This also causes these households to violate the fungibility assumption more often. However, this thesis focuses mostly on the difference in behavior between economists and non-economists. Economists are used to calculating when making decisions and generally have a good mathematical ability. Therefore economists are expected to engage less in mental accounting. This feeds the expectation that economists violate the fungibility assumption less often and are less sensitive to the labeling effect.

Not only labels that suggest a certain way to spend the income can influence consumer behavior. Just the name of the income source also matters for which mental account it will be diverted to. This is the conclusion of a paper by Epley, Mak & Idson (2006). The authors of this paper conduct several experiments and find that the name attached to the income influences the decision to save or consume this income. In the experiments subjects were surprised with an unexpected windfall of 50 euros. There were two conditions in which the subjects received this money: one group was told that it was a tuition rebate, where the other group was told that it was just bonus income. After receiving their check, subjects could leave. One week later they were asked how much of their money they had already spent. The result was that subjects in the bonus-condition spent significantly more (around $25 on average) than subjects in the rebate-condition (less than $10). The explanation for this difference can be found in the fact that a bonus was interpreted as a gain on top of the current wealth, where a tuition rebate was interpreted as a return to a previous wealth state. Therefore the money was more likely to be spent hedonistically in the situation where it was called a bonus. This shows once again that mental accounting matters and also that labels that do not recommend a specific way to spend the income can influence consumer behavior.

2.6 Summary

This section has presented a number of studies indicating that humans are much less rational in their choices than economic theory traditionally assumed them to be. In fact Dellavigna (2009) concludes that people systematically and consistently deviate from the rational choice model. They are sensitive to a lot of biases. The fungibility assumption, one of the main assumptions in the rational choice theory, does not always hold in a predictable manner. Several studies show that labeling can be a successful way of manipulating consumer choice. There is also evidence that not all population groups react in the same way to labeling. In the experiment by Abeler & Marklein it became clear that subjects with higher mathematical ability were less sensitive to the labeling effect. This finding provides the foundation for the expectation that economists, who generally have higher mathematical ability, will differ in their sensitivity to the labeling effect. To find out more about the differences in behavior between economists and non-economists in relation with the fungibility assumption and the labeling effect, an experiment has been conducted. Section 3 will present the main features of this experiment.

3. Experimental design

The goal of this thesis is to gather more information about the effects of labeling on consumption behavior in general, and on whether there is a difference in the sensitivity to labeling between economists and non-economists. In order to do this, a lab experiment has been conducted. In sections 3.1 and 3.2 the design of this experiment will be discussed in detail. Section 3.3 will discuss the expectations about subjects' behavior in the experiment. Finally, section 3.4 will discuss in more detail how the experiment was conducted.

3.1 Setup of the experiment

For this experiment several students of both economic and non-economic studies (at hbo-level or higher) were approached either by e-mail or in person at Tilburg University.[4] They were asked to participate in an experiment, in exchange for a chance on some money. This procedure should lead to a sample with both economists and non-economists. The group of economists has to be sufficiently large in order to be able to draw conclusions about their behavior. Therefore the sample will not be a representation of society, but just a means to do some tests on the difference in behavior of economists compared to non-economists.

During the experiment subjects were faced with four consumption decisions. For each of these four decisions subjects were given a budget that they could spend on two "goods". By doing so, they could earn "happiness points".[5] The goal of the experiment was the same for all subjects: they had to earn as much happiness points as possible. Each decision was framed in such a way that there was only one specific decision that maximized a subject's utility. To encourage active participation, I announced in the instructions that I would randomly select 3 participants out of all participants in the experiment after the experiment had finished.[6] Out of these 3 participants, the participant who earned the most points in this experiment would win 15 euros. The numbers 2 and 3 would win 10 and 5 euros respectively. As a consequence, earning more happiness points could actually lead to higher earnings for any subject. This mechanism was designed to give subjects a monetary incentive to participate actively and earn as many points as possible, to imitate the incentives people face in real life where they should also want to maximize their utility.

It was also mentioned in the instructions that the usage of a calculator was not allowed during the experiment. This was quite important, since with a calculator it would be too

[4] Note that Tilburg University specializes in social sciences and does not offer studies in nature sciences. Therefore the proportion of students studying technical studies in the sample will be quite small.
[5] These points can be seen as utility points, but in order to avoid confusion among subjects without a background in economics the name "happiness points" seemed more appropriate.
[6] A copy of these instructions can be found in Appendix A2.

easy for subjects to calculate the option that maximizes utility. Moreover, in real life consumers also go shopping without a calculator, so ensuring that all subjects would have to rely on nothing other than their own skills only adds more realism to the experiment.

In the first three treatments of the experiment subjects were given a fixed amount of virtual tokens as a budget. They could spend this budget (hypothetically of course) on two goods: drinks in a bar and study books. In the first two decisions they could freely spend their entire budget, where in the third treatment they had to spend part of their budget on study books, because they were faced with an in-kind benefit in that treatment, as will be discussed in section 3.2. In the first part of the instructions the buying procedure was explained in a clear way, illustrated by some examples. Subjects were also asked three example questions to make sure they really understood the experiment.

During the experiment the prices of the two goods remained equal to 2 and 4 tokens per unit and the budget was always equal to 40 or 60 tokens. This was done to keep the calculations simple for the subjects and to prevent subjects from miscalculating the amount of goods they could buy with their budget. To encourage subjects to spend their entire budget, unspent budget did not yield any happiness points. This would make it easier to compare consumption behavior between subjects.

After the experiment was finished, subjects were asked to answer some questions about their background. The answers to these questions were used for the analysis of the results. These background questions were related to gender, age, math skill, experience in economic experiments and field of study.

3.2 The four treatments

Treatment 1: the reference stage
The first decision subjects had to make (in the remainder of this thesis decisions will be referred to as treatments) was how to spend their budget of 40 tokens on drinks and study books, without any forms of manipulation. This treatment served as a reference stage for later decisions. In this treatment subjects were supposed to maximize the amount of happiness points they earned, according to a table that gave the amount of happiness points that they would earn if they bought that specific amount of goods. These tables are also included in appendix A2. Subjects were not shown any utility function, but the problems they were facing could be translated in a mathematical optimization problem, with the function

$$U_1(d,s)=24d^{0.7}+54s^{0.65} \quad (1)$$

as the utility function and their budget constraint would be

$2d+4s \leq 40$ (2)

In these functions, d stands for the amount of drinks consumed and s stands for the amount of study books bought by the subject. The budget constraint implies that subjects can buy at most 20 drinks or 10 study books, considering the prices of 2 and 4 tokens and the budget of 40. The values in this utility function were chosen at random, but in such a way that in each of the four treatments there was only one optimal decision. When solving the above mentioned optimization problem, we find that it would be optimal to purchase 8 drinks and 6 study books in this treatment.

Treatment 2: extra budget through wage rise
In treatment 2 subjects were first introduced with extra income in the form of a wage rise. They were told that a wage rise at their weekend job caused their budget to increase from 40 to 60 tokens. So the new budget constraint was

$2d+4s \leq 40$ (3)

Note that the extra income could be spent freely in this case. Subjects just received the extra income and could spend it freely over the two consumtion goods. So the only difference with the first treatment was that the budget increased. All prices and the values in the table remained the same. Therefore we have to solve a similar optimization problem as in treatment 1, the only difference being the increased budget. Solving this optimization problem, one would find that it would be optimal to purchase 12 drinks and 9 study books in treatment 2.

Treatment 3: extra budget through a labeled subsidy
In treatment 3 subjects again started with a budget of 40 tokens, but now they received a subsidy from the government to buy study books, worth 20 tokens (=5 study books). Therefore in this treatment the total budget is the same as in treatment 2: 60 tokens. The subsidy was given as an in-kind benefit, so the extra tokens had to be spent on study books. It is important to see that the optimization problem in treatment 3 is the same optimization problem as in treatment 2. Therefore subjects should purchase 12 drinks and 9 study books again to maximize their utility; this makes this treatment similar to the label treatment in Abeler & Marklein (2008). Since this optimal consumption bundle requires subjects to buy more than 5 study books anyway, the subsidy would not be distortionary and subjects did not have to change their optimal consumption bundle compared to treatment 2. If subjects did alter their consumption behavior, this would be a violation of the fungibility assumption.

Treatment 4: hours worked as choice variable
Treatment 4 was a bit different from the first 3 treatments. In this treatment subjects were not really faced with a consumption decision, but they had to work for two annoying clients in a nursing home and their goal was to minimize the happiness points they lost by doing that. This treatment was actually very similar to treatment one.

Subjects were explained they had a boss that required them work for 40 hours with two clients called Bassie and Adrian. So just like in treatment 1 they had an "endowment" of 40 hours (instead of 40 tokens) that they had to spend on two things: working for either Bassie or Adrian. Doing one job for Bassie would take 4 hours and one job for Adrian would take 2 hours. So in fact, subjects had to maximize their utility function, subject to a time constraint of 2a+4b=40, where a and b stand for one job done for Adrian and Bassie respectively. The only difference compared to treatment 1 was that subjects now had to minimize losses, where they had to maximize earnings in treatment 1. For rational subjects framing should be neutral, so they should not have any trouble finding the optimal distribution in treatment 4. If a subject acts less (or more) rational in this treatment, then a framing effect would be present. The optimal solution to this problem would be to do 4 jobs for Adrian and 8 jobs for Bassie, which yields a subject exactly the same amount of points as the optimal decision in treatment 1 (276). This will make it easier to compare a subject's performance in treatments 1 and 4 later.

An overview of all treatments is given in table 1:

Table 1: Overview of all four treatments

	Budget	**Budget constraint**	**Label on extra income**	**Framing**	**Rational decision**
Treatment 1	40 tokens	$2d + 4s \leq 40$	-	Gains	8 drinks + 6 study books
Treatment 2	60 tokens	$2d + 4s \leq 60$	Wage rise	Gains	12 drinks + 9 study books
Treatment 3	60 tokens	$2d + 4s \leq 60$, for $s \geq 5$	Book voucher	Gains	12 drinks + 9 study books
Treatment 4	40 hours	$2a + 4b = 40$	-	Losses	4 jobs for Adrian + 8 jobs for Bassie

3.3 Expectations

Rational choice theory obviously predicts that all subjects would act rational in this experiment. The main goal of this experiment is to show that labeling and framing can influence behavior. Therefore the expectation is that most subjects will act rational in treatment 1 and 2 and switch to irrational behavior in treatment 3 and 4 because of respectively the labeling effect and the framing effect. The exact hypotheses will be mentioned in section 4 when discussing the results; in this section it will be explained in detail why a labeling effect or a framing effect might occur.

First we will look into the labeling effect. As mentioned before, standard economic theory assumes that all subjects act as rational utility maximizers. Consider figure 1. In this figure the subsidized good (s) is on the horizontal axis and the non-subsidized good (o) is on the vertical axis. Subjects have a budget of R and the price of the subsidized good is normalized to 1. In the graph, point A denotes the optimal consumption bundle in the treatment without extra income or subsidy. When subjects are faced with extra income of size S (either through a wage rise or a subsidy by the government), their budget curve shifts to the right. In the case the extra income can be freely spent, subjects can choose any consumption bundle on the line between points D and F. A rational subject would then choose the optimal consumption bundle: point B in figure 1.

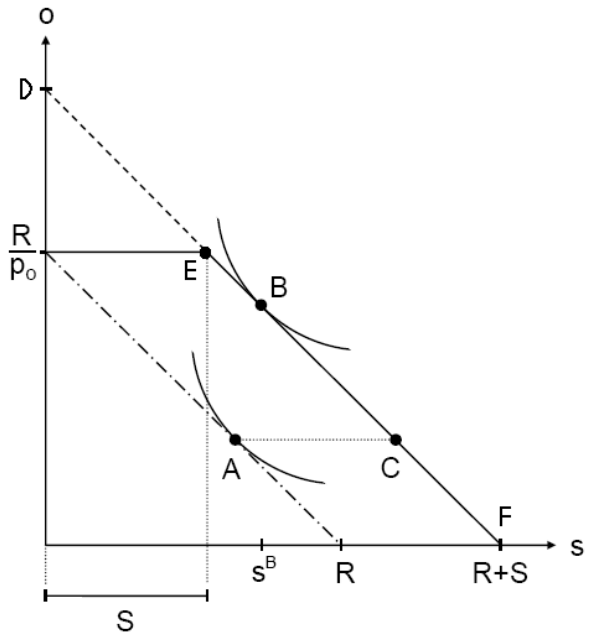

Figure 1: The optimal consumption bundle with and without subsidy (Source: Abeler & Marklein, 2008, slightly edited)

In the case that the extra income is an in-kind subsidy that has to be spent on the subsidized good (treatment 3) the budget constraint will be kinked. Subjects can then choose any consumption bundle on the line between points E and F. The subsidy is non-distortionary however and therefore a rational subject would still choose consumption bundle B. So under the assumptions of fungibility of money and full rationality, all subjects should choose consumption bundle B in both treatment 2 and treatment 3. It is likely however that some subjects will be sensitive to labeling and spend more than optimal on the subsidized good, choosing a consumption bundle somewhere between B and C.[7] Abeler & Marklein (2008) support this finding. In both their lab- and their field experiment they find that subjects spend more money on the subsidized good in the label treatment. Another conclusion by Abeler & Marklein is that this labeling effect is larger for subjects with lower mathematical abilities. In this thesis the difference between treatments 2 and 3 will be used to test whether a labeling effect is present and whether economists are less sensitive to this labeling effect compared to non-economists.

Tversky & Kahneman (1981) have shown that framing of choices can influence choice. They argued that it can make a difference whether the consequences of a choice are framed as gains or losses (see also section 2.2). In order to see whether such a framing effect is also present in consumer decision, the performances of subjects in treatment 1, where all consequences of their decisions were presented as a gain, will be compared to

[7] Note that consumption bundle C is the point where, compared to consumption bundle A, the entire subsidy is spent on the subsidized good.

their performances in treatment 4, where all consequences were presented as a loss. For a rational subject framing would not matter: in both treatments he or she would maximize utility and find the optimal allocation of their budget. If subjects are sensitive to framing however, they would perform better in either treatment 1 or treatment 4. The presence of a framing effect would indicate that subjects are better capable of dealing with either gains or losses. This knowledge could be useful for marketing campaigns by the government to discourage undesirable behavior. Consider for example the marketing campaign to discourage smoking. The government has two options at their disposal for such campaigns: they could either focus on the long and happy life without smoking or on the unhappy life full of diseases with smoking. If it would be the case that subjects are able to deal with gains better than with losses, then an implication of this result would be that the marketing campaign that focuses on the gains of non-smoking could be more effective than the current campaign that focuses more on the negative effects (losses) of smoking.

3.4 The actual experiment

The experiment was conducted during the last two weeks of May 2010. Potential subjects were approached through e-mail. In this e-mail they were asked to participate in an economic experiment (for the full text of this e-mail, see appendix A1). In the attachment they would find a document with the instructions of the experiment. Subjects were asked to read these instructions carefully and in this e-mail it was explained that it was not allowed to use a calculator during the experiment because otherwise their results would lose all of their value.[8] Since the response through e-mail was not sufficient to get a good sample, I also approached potential subjects in person at the food court of Tilburg University in the first days of June 2010. This is a food court outside the faculty buildings open to all students. These efforts finally resulted in a dataset containing the decisions of 52 subjects.

All of these subjects were students or recent graduates[9] at hbo-level or higher, in several different fields of study, so the subject group contains both economists and non-economists. It is also important to mention that almost all non-economists that participated in the experiment study various topics, both technical (for example chemistry or computer science) and non-technical (for example history or law). Therefore the group of non-economists is expected to be a reasonably representative sample of all higher educated students in non-economic studies. This will allow me to compare the differences in behavior between economists and non-economists.

[8] I am aware of the fact that I could never be sure that subjects did not use a calculator. However I think that stating that they were not allowed to use one both in the e-mail and in the instructions was the best solution out of the options I had at my disposal, since I was not allowed to conduct my experiment at the CentER-lab at Tilburg University.

[9] There is no reason to assume that these recent graduates act in a different way than students in their last year. Therefore these subjects were counted as students in their 5th year or higher.

Table 2 shows some sample statistics and aggregate results of the experiment. As can be seen in the table, around 56% of all subjects in the sample is a student in economics. This means that the sample is in no way representative for an entire population, but should be useful for the purpose of analyzing differences in decisions of economists and non-economists. Moreover, a majority of the sample is male (63%). This could also cause a bias, but it will be tested whether males and females act different by adding this dummy variable to the regressions. Furthermore, 38% gives themselves a 4 out of 5 or higher on the self-assessed variable mathematical ability. All these subjects will be referred to as the "high mathematical ability group". Also it was taken into account whether the subject indicated to have some experience in economic experiments and whether the subject submitted his/her e-mailaddress. These variables will not be used in the analysis, but will be useful later on. A first indication from this table that the rational choice theory may not hold for a lot of subjects is the observation is that there is a lot of variation in the number of goods they bought. This indicates that at least some subjects did not act in line with the rational choice model.

Table 2: Results and statistics.

		Optimal	Average	St. dev.	Minimum	Maximum
T1	Amount of drinks purchased	8	7.08	3.49	0	18
	Amount of study books purchased	6	6.46	1.74	1	10
T2	Amount of drinks purchased	12	11.27	4.30	0	28
	Amount of study books purchased	9	9.37	2.15	1	15
T3	Amount of drinks purchased	12	9.85	4.37	0	20
	Amount of study books purchased	9	10.08	2.19	5	15
T4	Amount of jobs done for Adrian	4	5.88	4.94	0	20
	Amount of jobs done for Bassie	8	7.06	2.47	0	10
	1 if economist		0.56		0	1
	1 if female		0.37		0	1
	Age		21.85		18	29
	Mathematical ability (self-assessed)		3.09		1	5
	Experience in economic experiments (self-assessed)		1.96		1	4
	Year		2.69		1	5
	1 if high mathematical ability		0.38		0	1
	1 if e-mailaddress submitted		0.79		0	1
	1 if experience in economic experiments		0.54		0	1
	Percentage of points scored		91.26		47.86	100

Note: A definition of all the variables used can be found in appendix A3.

The next section will discuss the results of the experiment.

4. Results

This section will report the results of the experiment, by means of testing several hypotheses. These hypotheses will be presented as a null-hypothesis that assumes the rational choice theory holds. In case this null-hypothesis can be rejected, this would mean a violation of the rational choice theory. First it will be tested whether economists act more rational than non-economists. After that, the presence of a labeling effect will be discussed, as well as the differences in sensitivity to this labeling effect between economists and non-economists. Finally it will be tested whether it makes a difference if the consequences of choices are framed as gains or losses. Section 4.5 will then discuss the robustness of the results.

4.1 Basic results

Figure 2 gives an overview of the decisions made in treatment 1. In this treatment it would have been optimal to purchase 6 units of the subsidized good: study books.[10] Around 45% of all subjects that participated in this experiment made this optimal decision in this treatment, which was similar to the reference stage in the experiment by Abeler & Marklein. In their experiment the percentage of subjects that made the optimal decision in this reference stage was slightly lower: around 38%. This difference can be caused by several factors, some of which can be considered as a weakness in their research:

- The dataset by Abeler & Marklein did not include economists. This could mean that subjects in their experiment were less calculating and performed worse in the experiment. This will be the first hypothesis that will be tested.
- In this experiment subjects had more incentives to calculate: if subjects earned 1 point less in the experiment, this could cost them 5 euros. In the experiment by Abeler & Marklein each point was worth 0.01 euro. The incentives for a subject to put in a lot of effort to earn a few cents more were probably not that high and therefore subjects were more likely to settle for an acceptable option rather than searching for the best option in the experiment by Abeler & Marklein.
- In the experiment by Abeler & Marklein prices were equal to 2 and 3 and the budget was equal to 50. This made calculations harder for subjects and therefore they could have been less likely to find the optimal consumption bundle.
- The more comfortable environment in which my experiment was filled in could also have played a role here. More than half of all subjects were able to fill in the experiment at home. Abeler & Marklein did their experiment using Z-Tree software, where a group of subjects participates simultaneously. Even if the researchers did not add a time limit to the experiment, subjects that needed

[10] Note that there was no subsidy present yet in treatment 1, but to make the analysis more general and easy to compare with existing research, I will from now on refer to study books and drinks as the subsized good and the non-subsidized good respectively.

time to calculate the optimal consumption bundle could be discouraged to make these calculations by the fact that they kept all other subjects waiting. As a consequence of this, some subjects in their experiment could not have found the optimal solution.

Since the most likely explanation for this difference is the fact that economists were included in the dataset (for evidence: see section 4.2), this experiment was probably quite similar to the experiment by Abeler & Marklein. This means we can compare the results from this experiment to the results from Abeler & Marklein and test whether their conclusions also hold when economists are included in the sample. The next section will discuss whether economists act significantly different from non-economists when it comes to rationality.

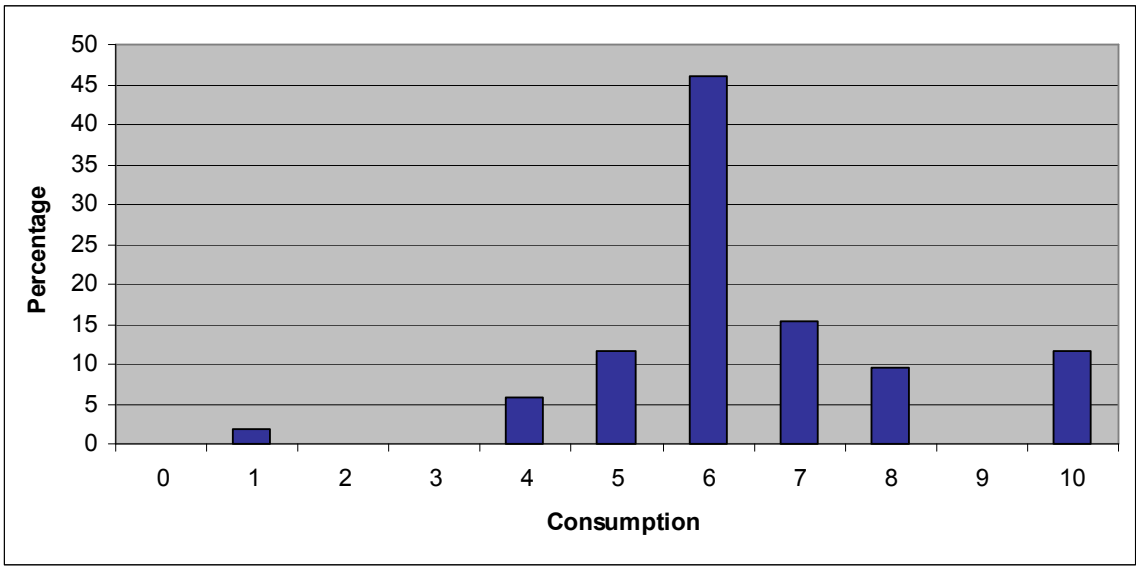

Figure 2: Consumption of the subsidized good in treatment 1. Note that it was optimal to buy 6 units of the subsidized good in this treatment.

4.2 Rationality and education

The previous section has given evidence that subjects did not always act in line with the rational choice model. This is in line with findings in the behavioral economics literature, as summarized in section 2 of this thesis. The present section will look into the differences in rationality between economists and non-economists throughout the whole experiment. A completely rational subject would choose the optimal consumption bundles as mentioned in tables 1 and 2. The results in table 2 indicate that a considerable amount of subjects did not act rational throughout the whole experiment. In fact, only 14 out of the 52 subjects made the optimal decisions in all 4 treatments and scored the maximum score of 100%. Out of these subjects, 11 were economists and 3 were non-economists, all of which did technical studies.[11] This finding seems to support

[11] To be precise, these three subjects studied chemistry, computer science and industrial engineering.

the hypothesis that economists act more rational than non-economists. Economists are much more used to making calculations when confronted with numbers and are trained at school to maximize utility. Therefore it seems likely that these subjects perform better in this experiment.

Hypothesis 1: Economists act just as rational and calculating as non-economists.

Table 3 provides the results for the OLS-regression with the total percentage of the points scored as dependent variable. The estimates in column 1 show that without controls the dummy that indicates whether a subject is an economist is significant at 5%-level (p-value = 0.025). When control variables are added to the regression, the significance of the dummy does not change much (p=0.023), but the coefficient does increase a bit. The estimate for this dummy in column 2 indicates that economists on average score 10.5% more points than non-economists, so this means that economists act significantly more rational and calculating in this experiment. Therefore hypothesis 1 can be rejected and we have found a violation of the rational choice model.

Another conclusion that can be drawn from table 3 is that subjects with high mathematical ability perform significantly better than other subjects (p=0.027). Since subjects were not allowed to use a calculator and had to perform some calculations in order to find out which consumption bundle yielded the higher payoff, this is not a surprising result. The other explanatory variables in column 2 are not significant, as expected.

Hypothesis 2: Economists do not increase in their rationality as they approach the final stages of their study.

In column 3 of table 3, the coefficient of the variable year*"1 if econ" provides an estimate for the effect of one additional year of economic education on performance in the experiment.[12] The negative coefficient seems to indicate that economists perform worse as they get further in their study, which is a surprising result. The result is far from significant though, so we cannot conclude any increase or decrease in rationality based on this coefficient. Hypothesis 2 can therefore not be rejected, so there seems to be no real effect of economic education on rationality and thus the reason why economists act more rational than non-economists must be found in other factors, such as personal (possibly innate) qualities. Based on these results, it is also not possible to reject a hypothesis that the difference between economists and non-economists is created during the first year of economic education. This would require some additional research.

[12] Note that the dummy that indicates whether a subject is an economist is not included in the regression, since the effect of economic education is now included in the two new variables..

Table 3: OLS-estimates for factors influencing performance in the experiment and the effect of labeling on consumption.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	**Percentage of points scored**			**Treatment-difference**		
1 if economist	9.143**	10.490**		-0.829**	-1.121**	
	(3.968)	(4.465)		(0.402)	(0.468)	
1 if high mathematical ability		9.672**	11.279**		-0.052	-0.198
		(4.229)	(4.245)		(0.443)	(0.457)
Age-18		1.198	1.161		-0.164*	-0.140
		(0.923)	(0.922)		(0.097)	(0.099)
1 if female		3.711	2.221		-0.400	-0.186
		(4.837)	(4.699)		(0.507)	(0.506)
Year		-2.164			0.127	
		(1.677)			(0.176)	
Year * "1 if econ"			-0.802			-0.016
			(1.595)			(0.172)
Year * (1-"1 if econ")			-4.132*			0.245
			(2.121)			(0.228)
Constant	86.161***	81.551***	87.070***	1.174***	1.795***	1.181*
	(2.963)	(6.255)	(5.665)	(0.301)	(0.655)	(0.610)
N	52	52	52	52	52	52
Adjusted R-squared	0.08	0.12	0.11	0.06	0.04	-0.02

Notes: standard errors are in parentheses; the dependent variables are the percentage of points scored in the experiment and the difference in consumption of the subsidized good in treatments 2 and 3 (for an exact definition: see appendix A3). Significance at 1%, 5%, and 10% level is denoted by *, **, and *, respectively.**

An interesting result of this regression though is that the coefficient for the variable year*(1-"1 if econ") now becomes negative and significant at 10%-level (p=0.058). This seems to indicate that subjects earn fewer points per year of non-economic schooling. A potential explanation for this result could be that students of non-economic studies become less used to calculating the further they get into their study. Therefore non-economists in further stages of their study could be more tempted to intuitively make their decisions without making any calculations. This could explain why they score fewer points than economists. However, this effect is only significant at 10%-level and therefore more research needs to be done on this subject to find out whether this effect really exists or if it is just coincidence and caused by this relatively small sample.

4.3 The labeling effect

Based on the results in section 4.2, it seems that the expectation that economists behave differently from non-economists in this experiment can be confirmed. However, the main goal of the experiment was to find out more about the response of economists to the labeling effect. First it will be tested whether a labeling effect is present in the dataset. As mentioned before, a rational subject would choose the same consumption bundle in treatments 2 and 3. If a subject would be sensitive to the labeling effect, he or

she would spend more on the subsidized good in treatment 3. This subject would then not treat income as fungible: a large proportion of the subsidy for study books would actually be spent on the subsidized good. In other words: the marginal propensity to consume the subsidized good would not be the same in treatment 2 and 3.

Table 2 gives a first hint that a labeling effect might be present: in treatment 2 subjects buy on average 9.365 units of the subsidized good. In treatment 3, the average consumption of the subsidized good is 10.077, which is 0.712 units higher than in treatment 2.

Hypothesis 3: A labeling effect is not present; consumption of the subsidized good is equal in treatment 2 and treatment 3.[13]

Figure 3 shows the change in consumption of the subsidized good in treatment 3 compared to treatment 2. A majority of all subjects keeps their consumption bundle (optimal or suboptimal) unchanged in treatment 3. However, 19 out of all 52 subjects chose different consumption bundles in treatments 2 and 3. These subjects violate the fungibility assumption. Since almost all of these subjects consume more of the subsidized good in treatment 3 (see figure 3), this violation is probably caused by a labeling effect.

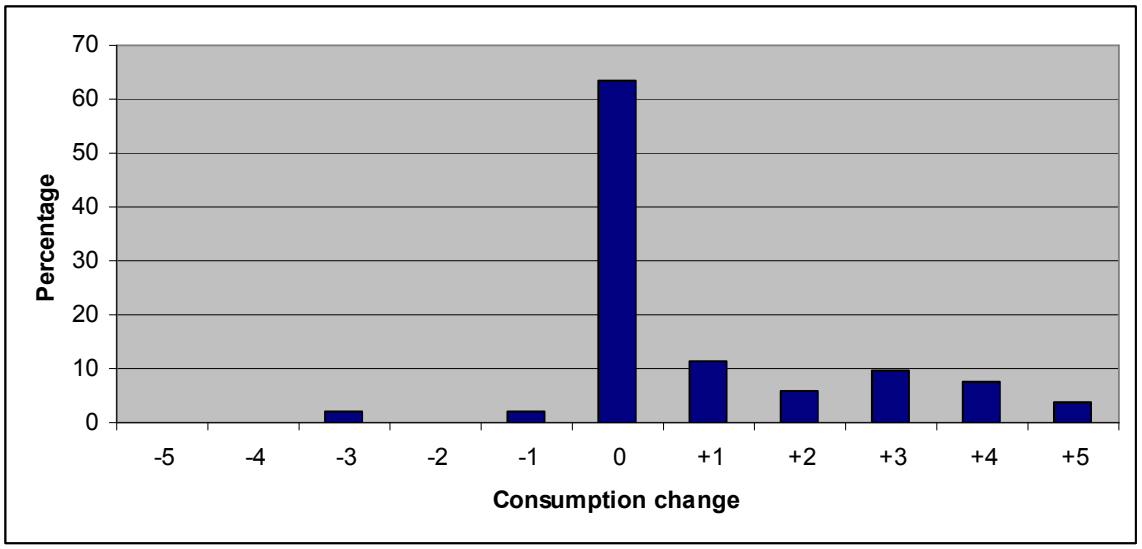

Figure 3: Consumption change of the subsidized good in treatment 3 relative to treatment 2. Note that it was optimal to have the same consumption bundle in both treatments.

[13] The most preferred way to test this hypothesis would be to randomly assign subjects to either the label treatment or the treatment where the extra income was not labeled and then create a dummy to test whether subjects purchase more units of the subsidized good in the label treatment. However, since I did not get the opportunity to do my experiment in the CentER-lab at Tilburg University, I had to make subjects participate in both treatments in order to get enough observations. This impedes my possibilities to estimate the labeling effect, but it is better than the alternative of not having enough observations to get any significant results.

Now let us look at the marginal propensity to consume the subsidized good out of the subsidy (MPC).[14] A rational individual would increase consumption of the subsidized good from 6 units in treatment 1, to 9 units in treatments 2 and 3. The costs of purchasing these 3 additional goods would be equal to 12 tokens out of the subsidy of 20 tokens. This means that the optimal MPC is equal to 0.6. When we look at the dataset, we find that in treatment 2 the average MPC is equal to 0.58. This is quite close to 0.6, so we can conclude that on average subjects act quite rational in treatment 2. If a labeling effect would be present, the MPC in treatment 3 would be significantly higher than the MPC in treatment 2. This seems to be the case: in treatment 3 the marginal propensity to consume out of the subsidy is 0.72. To statistically test the presence of this labeling effect, a t-test for two paired samples will be used to test whether the MPC in treatment 2 is equal to the MPC in treatment 3. With a one-sided test (since we expect the difference between the two MPC's to be larger than zero) the hypothesis that the two means are equal can be rejected in a very convincing way. (p=0.0006) Therefore based on this t-test we can conclude that a labeling effect is probably present and that hypothesis 3 can be rejected. This is another violation of the rational choice theory.

To prove the robustness of this conclusion, the same conclusion will be obtained by a Wilcoxon matched-pairs signed-ranks test. Since this is not a very common test, the test procedure is explained in appendix A4. When conducting this test, we find that $W_+=172.5$ and $W_-=17.5$, so $W=17.5$. There are 19 observations, so we have to compare this value to the critical value of 32. The test statistic is clearly smaller than 32, so the conclusion of this test is that the difference between the two MPC's is significant at 0.5%-level with a one-sided test. This supports the conclusion from the t-test that a labeling effect is present. Now it would be interesting to test whether economists are less sensitive to the labeling effect.

Hypothesis 4: Economists are equally sensitive to the labeling effect as non-economists.

The alternative If economists would be less sensitive to the labeling effect, this would mean that the difference in their consumption of the subsidized good in treatments 2 and 3 would be smaller. Column 4 of table 3 shows the results of the regression with the difference in the consumption of the subsidized good (see figure 3) as dependent variable and the dummy that indicates whether a subject is economist as explanatory variable. The dummy is significant (p=0.045) and has the expected sign. When adding more control variables, the dummy becomes even more significant (p=0.021), but all of the control variables are insignificant, including mathematical ability. This confirms the

[14] Note that this MPC would be equal to 1 if the entire subsidy would be spent on the subsidized good, and that it would be equal to 0 if the entire subsidy would be spent on the non-subsidized good.

expectation that economists are less sensitive to the labeling effect than non-economists.[15]

Now it would be interesting to see if economists become less sensitive to the labeling effect near the end of their study. In column 6 of table 3 the variable year*"1 if econ" is added to the set of explanatory variables again. However, this regression does not give any significant coefficients at 5%-level. Therefore the hypothesis that economists become less sensitive to the labeling effect as they get further in their study has to be rejected. This could mean that the explanation for the fact that economists are less sensitive to labeling does not lie in the fact that economists are trained to see through labeling, but again in their personal qualities (or the first year of economic education).

Furthermore columns 5 and 6 in table 3 give evidence that there is no significant difference in the sensitivity to the labeling effect between males and females in this experiment. This contradicts the conclusion by Lyssiotou (2009, see section 2.4) that females are more sensitive to the labeling effect. However, it is likely that the gender difference as shown in the paper by Lyssiotou is only present in special cases. In these cases the difference can most likely be explained by other factors, such as altruism. Altruism was not present in this experiment, but it is certainly present when it comes to deciding on how to spend child benefits.

4.4 The framing effect

This section will discuss whether a framing effect is present. In other words, we will test whether it makes a difference for a subject's performance if earnings in a similar treatment are framed as gains or losses.

Hypothesis 5: Subjects in the experiment were not sensitive to gains/losses framing.

Obviously for a rational subject framing should not matter. In order to test this hypothesis, two new dummies have been created, indicating whether a subject made the optimal decision in treatment 1 and 4 respectively. In treatment 1 46.15% of all subjects chose the optimal consumption bundle and in treatment 4 this percentage was 42.31%. A t-test for two paired samples has been conducted to test whether the mean of these two variables is different. The difference is not significant however ($p=0.5323$, two-tailed test) and therefore we can conclude that framing did not significantly change the amount of rational choices.

To see whether there is an effect of framing in the relative performance of subjects compared to the maximum amount of points they could earn (which was equal to 276 in

[15] This conclusion also holds when we use a t-test. When using a sample of only economists, the average MPC's in treatment 2 (0.61) and treatment 3 (0.68) are not significantly different any more. This seems to support the conclusion that economists are less sensitive than non-economists to the labeling effect.

both treatments), it will be tested whether subjects scored more points in either treatment 1 or treatment 4. However, the average number of points scored in these two treatments differs only very little: the p-value of the t-test testing whether these two means are different is 0.7166. Therefore we cannot conclude that subjects in this experiment act different when facing losses instead of gains, so the hypothesis that subjects are not sensitive to gains/losses framing in this experiment cannot be rejected.

A potential explanation for this could be that subjects still earned a positive amount of points by making their decision. Furthermore the points that could be earned in treatment 4 were not directly linked to a certain amount of money. As a consequence of this, decision 4 was seen by a subject as just another opportunity to earn points and therefore loss-aversion, which caused the framing effect in the experiments by Tversky & Kahneman (1981), played a very limited role in this experiment or maybe even no role at all because subjects did not interpret the conseqeunces of treatment 4 as losses.

4.5 Robustness of results

When subjects are participating in an economic experiment, a lot of factors can influence their performance. Some of the most obvious factors like age, gender and study have been controlled for in the analysis of the results. However, there could still be some hidden factors that influence the performance or motivation of subjects. This section will prove that the conclusions of this thesis are robust to the most important factors.

As mentioned in the instructions, a subject's understanding of the experiment is very important. In the instructions three example questions had to be answered and subjects could either check these answers themselves (in the case they participated at home) or had them checked by the experimenter (in the case they participated in the food court of Tilburg University). In both cases the subject got an explanation of the answers, if necessary. However, it could still be that some subjects did not understand the experiment properly. In order to test whether the subjects that did not understand the experiment influenced the results, a new dataset was constructed excluding all subjects that chose to consume 2 or less units per product in one or more of the first three treatments.[16] This procedure eliminates 7 observations from the sample. When we redo regressions (2) and (5) without these 7 observations (see columns (i) and (iv) in appendix A5), we find that the main conclusions are robust to the exclusion of subjects that might not have understood the experiment properly. In fact, the significance of the coefficients of interest increases a little. However, the coefficients themselves decrease

[16] These decisions *could* indicate that the subject did not understand the experiment, but this is not certain, since the observation that a subject that decides to buy 0 drinks and 15 study books does not mean the subject did not understand the experiment: he or she could also have made the decision intuitively and without any calculations. This would then still be a meaningful observation. Therefore these observations cannot be removed from the actual dataset and will only be removed in this section to test the robustness of the conclusions based on the entire dataset.

a little, but this is caused by the increase in the constant due to the exclusion of the worst performing subjects. Since the significance of the coefficients of interest increases, this does not change the general results.

Also the labeling effect is robust to the exclusion of these 7 observations, although it does become a bit less significant. The p-value of the one-sided t-test increases to 0.003 (where it was 0.0006 with the inclusion of these observations), but the difference in the MPC between treatment 2 and 3 remains highly significant.

It could also be that differences between economists and non-economists arise due to the fact that economists have more experience in economic experiments. However, doing regressions (2) and (5) again with an extra dummy variable indicating whether a subject has at least some experience in economic experiments does not change much to the coefficients of interest and their significance (see columns (ii) and (v) in appendix A5). Moreover, the dummy variable itself is very insignificant. The labeling effect is also robust to excluding observations that did not have any previous experience in economic experiments, but the p-value increases to 0.020 for a one-sided test. This means that the labeling effect is not significant at 1%-level any more. This increase can partly be explained by the smaller sample size (only 28 observations had some experience in economic experiments), but it could be an indication that subjects with some experience in economic experiments are a bit less sensitive to the labeling effect. However, the most likely explanation for the increase in the p-value would be that 21 out of all 28 observations are economist in this sample. This increase in the p-value seems to be in line with the conclusion that economists are less sensitive to the labeling effect than non-economists.

It could be the case that some subjects do not do this experiment for money and therefore are less calculating, because they know in the end they will not earn anything with a high amount of happiness points anyway. In such a case a subject could be much more likely to randomly select a consumption bundle and therefore performs worse than subjects that are able to earn money with their decisions. Columns (iii) and (vi) in table A5 however show that there is no reason to assume that subjects who did not fill in their e-mail address (so subjects that did not have a chance to win any money) perform worse than subjects that did fill in their e-mail address. The dummy indicating whether a subject entered an e-mail address is very insignificant in both regressions and all conclusions still hold. Moreover, when removing all observations of subjects that did not enter their e-mail address from the sample, the labeling effect is still significant using the standard t-test. (p=0.004 for a one-sided test)

Therefore all the conclusions as presented before are robust to all these factors that have the potential to influence a subject's performance.

4.6 Conclusion

This section has presented the results of the experiment. It has been shown that a labeling effect was present in the experiment: 19 out of all 52 subjects violated the fungibility assumption by changing their consumption bundle in treatment 3 compared to treatment 2, even though nothing had changed except for the label on their extra income. Furthermore this section has shown convincing evidence for the hypotheses that economists act more rational than non-economists and are less sensitive to the labeling effect. The explanation for these findings can however not be found in their education, but possibly in the personal (innate) qualities of economists. The next section will discuss the potential implications of these results for policy makers.

5. Policy implications

In the previous section we have found significant differences in behavior between economists and non-economists and strong evidence for a labeling effect. In this section the implications of these results will be discussed. A government should be very interested in these results. If labeling can influence choice behavior, this means that a government should take into account that this mechanism could be used to help its citizens make better choices, but also that third parties may use it, which could harm consumers. A government should be very aware of this and take action when necessary.

The finding that economists and non-economists differ in their sensitivity to the labeling effect is also interesting. In order to use labeling in an optimal way, it is necessary to fully understand what effects labeling could have on choice behavior and which groups are the most effected by this mechanism. Section 5.1 will compare the results of the experiment with existing literature; sections 5.2 and 5.3 will discuss the implications of the results in the experiment for a government; section 5.4 will discuss the role of the government as choice architect and section 5.5 will provide an overall conclusion.

5.1 Discussion of results

The results of the experiment show, consistent with Kooreman (2000) and Abeler & Marklein (2008), that a label attached to income can influence consumer behavior. This means that the conclusion by Abeler & Marklein is robust to the inclusion of economists in the sample. However, there are some differences with their results. In contrast to what they concluded, with this dataset I found no effect of mathematical ability on the sensitivity to the labeling effect.[17] This is a surprising result that cannot easily be explained.

The results also seem to indicate that some individuals satisfice utility instead of maximizing it (see section 2.1). Figure 2 showed that individuals often made a suboptimal but close to optimal decision in treatment 1. Moreover, in treatment 2 and 3 the easy (but suboptimal) consumption bundle consisting of 10 units of each good was quite common, as around 20% of all subjects chose this bundle in both treatment 2 and 3. This is an interesting result, which shows that some individuals rather pick the easiest choice than the best choice.

Another result of the experiment is that many people do not act rational. The explanation for this observation can be found in the fact that scientific thinking goes against our nature. Earlier the question was raised whether scientific thinking can be taught. The results of this experiment imply that there is no effect of economic education itself on rationality or sensitivity to the labeling effect. This seems to be

[17] This conclusion also holds when removing the dummy "1 if econ" from the set of explanatory variables, although the dummy indicating high mathematical ability becomes a bit less insignificant.

evidence against the conclusion by Henk Schmidt (see section 2.4) that scientific thinking can be taught, since one would think economists should be one of the first groups that will be trained to maximize utility, based on their education.

However, the results do seem to indicate that economists act more rational than non-economists. This raises the question whether rationality is innate and cannot be taught at all, or if it is the first year of economic education that makes economists act more rational than non-economists. Based on this experiment, this question cannot be answered yet. Another important observation is that economists were less sensitive to the labeling effect. The same seems to be true for students of technical studies, but from this dataset it cannot be concluded that these students perform significantly better than other students, because there are only 3 observations on this matter. These observations do however support the statement that some population groups are more rational and less sensitive to the labeling effect.

Finally, based on this experiment we cannot conclude that a framing effect is always present. The reason for this was probably that the sentiment of loss-aversion that normally causes the framing effect was not that strong in treatment 4. This does not undermine the conclusion by Tversky & Kahneman (1981), but it just adds some more insight into when loss-aversion is likely to influence consumer choice and when it is not.

To summarize, the most important observations from the experiment that will be used in the rest of this section are:

1. A labeling effect is present
2. There are certain groups that are less sensitive to the labeling effect
3. The groups that are less sensitive to the labeling effect are also the groups that acted the most rational in this experiment

5.2 Motives for government intervention

As shown in the previous sections, humans are sensitive to several biases that cause them to deviate from the path that maximizes their utility. Obviously this leads to a welfare loss for society. A paternalistic government can intervene in such a case and help individuals reach a higher level of well-being. Another reason why a government may want to interfere with individual behavior is the presence of negative externalities that are often not taken into account when making a consumer decision. Examples of negative externalities the government is currently trying to deal with are smoking (imposed on the people around you) or eating unhealthy (imposed on your future self). It is then important for a government to know what mechanisms could be successful in influencing consumer choice and know more about the costs and benefits (effectiveness) of these mechanisms.

Traditional paternalistic regulation only aims to protect individuals that are not capable to make good decisions themselves, such as children and adults in financial distress

(Kooreman & Prast, 2010). However, nowadays people have to make a lot of complex decisions, while being confronted with too much information to process. As a consequence, also competent people can make predictable errors in decision making. This increases the scope of paternalistic regulation: there are more groups that benefit. Traditional paternalistic policies include laws, fiscal incentives, providing information and educating people through warnings or marketing campaigns. These policies are more or less based on the rational choice model. A government seems to think that as long as people know what is good for them, they will behave accordingly. Just providing information proves to be not a very effective way of changing behavior. Moreover, laws impede free choice, taxes and subsidies create distortions and marketing campaigns are usually costly and not very effective, as it is just another way of providing information. Traditional paternalism therefore seems to be not very succesful.

Recent developments in the field of behavioral economics provide new foundations for paternalism. The use of defaults, either to encourage desired behavior (Choi, Laibson, Madrian & Metrick, 2003) or to discourage undesired behavior (Cronquist & Thaler, 2004), is proven to be a successful and cost-effective mechanism to alter choice behavior. As discussed in section 2, also the use of framing and labeling has the potential to influence behavior. A government can then play the role of choice architect: they can change the environment in which choices are made, which in turn affects individual choice. Choice architecture can nudge people to make better choices, without limiting freedom of choice (Thaler, Sunstein & Balz, 2010). This is called libertarian paternalism. Libertarian paternalists try to steer individual behavior in welfare-enhancing directions while maintaining freedom of choice. Thaler & Sunstein (2003) defend libertarian paternalism stating that people's preferences are often unstable and ill-formed. Under these circumstances, a government should do everything it can to avoid random effects and try to create a situation that increases welfare compared to the situation where the government would abstain from intervention. Libertarian paternalism would be a good instrument to do this, because it would not impede freedom of choice for consumers and could benefit some individuals.

5.3 Asymmetric paternalism

Some economists do not like the idea of a paternalistic government. Section 5.2 has taken the edge off the argument that paternalism limits freedom of choice, but there are still some critics that state too much paternalism can harm those who do not need government intervention. This critique is however mostly based on traditional policy tools such as taxes, subsidies and legal prohibitions and obligations. New policy tools such as labeling and other forms of framing, when used in the correct way, do not have distortions as large as some of the traditional policy tools. As section 4 shows, some population groups are less sensitive to labeling than other groups. Another important observation is that the groups that are less sensitive to labeling are in general more rational than other groups. Camerer, Issacharoff, Loewenstein, O'Donoghue & Rabin (2003) distinguish two kinds of consumers in their paper: fully rational consumers and

boundedly rational consumers. They call a policy asymmetrically paternalistic if "it creates large benefits for those people who are boundedly rational while imposing little or no harm on those that are fully rational" (p. 1219). Setting good defaults is a well-known way of asymmetric paternalism (see for example Choi et al., 2003), but also requiring firms to mention seemingly unimportant information, such as the total costs of purchasing a product and paying for it later in fixed payments, could benefit certain consumers while leaving fully rational consumers unaffected.

Also labels can play an important role in asymmetrically paternalistic policy. In the experiment labeling was not used for a paternalistic purpose[18], but obviously labeling can be used to increase well-being as well. Think for example of the logos on products in the Dutch supermarkets promoting a healthy choice. This kind of labeling will only influence behavior of ignorant consumers that do want to eat healthy, but do not because of their lack of information about the healthiness of their food. An argument against this kind of paternalism is proposed by Glaeser (2006), who states that soft paternalism can be seen as an emotional tax, in this case, to those that do not obey to the healthy food labels. This argument may be valid, but the question is what is more important: the health improvement of citizens that are sensitive to the labeling effect or the emotional tax on those unwilling to change their behavior. As observed by Camerer et al. (2003:1254) a few years earlier: "(asymmetric paternalism) has the potential to shift the debate from one about whether or not paternalism is justified, to one about whether the benefits of mistake prevention are larger than the harms imposed on rational people."

5.4 Choice architecture

In this thesis it has become clear that humans do not always make decisions in a vacuum-like environment. The environment has many features that can, noticed or unnoticed, influence the choices people make. The government can partly control this environment and be a choice architect to nudge people to make better choices. An important role of a paternalistic government is to correct so-called "expect errors" (Thaler et al., 2010). These are errors that are common to humans and can therefore be expected to occur more often. A government should not base its policies on the assumption that all individuals are rational, but focus on the expect errors that their citizens make. Therefore, a well-designed policy expects citizens to make errors and tries to help them prevent or correct these errors.

Another role for the government is to provide information in a stimulating environment. Providing information through national marketing campaigns may be effective, but it is also expensive. However, as shown in section 2, consumers do lack information without government intervention and may suffer from this (Krosnick, 1991). Therefore a

[18] In fact, subjects that were sensitive to the labeling effect generally ended up worse off in treatment 3 (the labeling treatment).

government with paternalistic motives would want to intervene and either supply information themselves or lower information costs. In the experiment it was clear that subjects with low mathematical ability scored significantly less points than subjects with high mathematical ability. In the experiment there were no real information costs, but finding the optimal consumption bundles did take some calculations. Therefore it could have been that the information costs observed by subjects with low mathematical ability were higher, causing them to satisfice their utility instead of maximizing it. This could explain the pattern in figure 2: a lot of subjects purchase between 4 and 8 units of the subsidized good in treatment 1, because they knew the utility-maximizing option was somewhere in this area, but the observed costs of calculating the exact utility-maximizing option were too high.[19] A government could use the finding that subjects in this experiment as well as normal individuals do not like to make too many calculations very well. For example in the credit-card market, consumers tend to underestimate the compounding effect of interest (Laibson, Repetto & Tobacman, 2001). Requiring firms to mention accessible information about the compounding effect of interest could benefit people that are potential victims of high credit card debts, while leaving rational individuals unharmed: another great example of asymmetric paternalism.

The same conclusion holds for savings for retirement, as shown by Card & Ransom (2006), and health-related behavior, as shown by Downs et al. (2009). In these cases the use of labels can be a succesful mechanism to encourage the desired behavior and increase the well-being of citizens. This shows that labeling is a successful way for a government to influence behavior at relatively low costs and provide new foundations for libertarian (and asymmetric) paternalism.

5.5 Conclusion

Knowledge from the field of behavioral economics can be very useful for a government when designing policies. Setting good defaults, framing and labeling can be a successful way of making citizens make better decisions while maintaining freedom of choice. This is called libertarian paternalism. The experiment has shown that not all population groups are equally sensitive to government manipulation and that one of the least affected groups, economists, was also the most rational group. This implies that a policy like labeling could benefit irrational people, while doing little or no harm to rational individuals (Camerer et al., 2003). This is called asymmetric paternalism. Based on these considerations, one could say that recent research in the field of behavioral economics has provided new foundations for paternalistic policies.

[19] The distributions of choices in treatments 2, 3 and 4 look similar to figure 2, so this conclusion holds for all treatments.

6. Conclusion

This thesis has provided convincing evidence for a labeling effect as presented before in the paper by Abeler & Marklein (2008). Moreover, the presence of this labeling effect is robust to the inclusion of economists in the sample, a conclusion that could not be drawn from the paper by Abeler & Marklein. However, it must be mentioned that economists appear to be less sensitive to labeling than non-economists. Furthermore there is some evidence indicating that economists are more rational compared to non-economists. This is an interesting conclusion, because it shows that labeling can be beneficial for those that are boundedly rational while doing little or no harm to those that are fully rational. Therefore labeling can be part of asymmetrically paternalistic policy.

Labeling can be useful in several domains, such as health-related behavior and finance to encourage the desired behavior. Labeling is a good mechanism to do this, because the costs are low, the effectiveness is high and it does little or no harm to individuals that are fully rational already. Therefore freedom of choice is maintained.

The hypothesis that subjects acted differently in the experiment when facing losses instead of gains could not be confirmed based on the results of the experiment. A potential explanation for this result would be that subjects did not consider the consequences of their final decision as actual losses and therefore the expected loss aversion that would influence behavior did not occur. Based on these results, nothing can be said about the effectiveness of this kind of framing as a mechanism to influence consumer behavior.

An interesting topic for further research could be to test whether the conclusions of this master thesis also hold for subjects that are not students. Is an individual that has studied economics several years ago still less sensitive to the labeling effect or does the effect of economic education fade as time passes? Another potential extension to this research would be to include lower-educated individuals in the sample. This sample included only students at hbo-level or higher in the sample and the sample of Abeler & Marklein included only students from the University of Bonn. It would be interesting to test whether education level also has an effect on the sensitivity to the labeling effect.

This master thesis has once again proven the existence of a labeling effect by means of an experiment. An important observation from this experiment was that certain groups seem to be less sensitive to the labeling effect. This shows that labeling can be part of asymmetrically paternalistic policy. Such a policy helps some individuals make better choices, while leaving others unharmed. This eliminates most traditional arguments against paternalism. Therefore asymmetric paternalism could very well be the future of paternalism, which would mean that we will see a lot more labeling in the future.

7. References

Abeler, Johannes and Felix Marklein (2008), *Fungibility, Labels, and Consumption.* IZA Working Paper no. 3500

Beshears, John, James Choi, David Laibson, Bridgitte Madrian (2008), *How Are Preferences Revealed?* Forthcoming in Journal of Public Economics.

Blow, Laura, Ian Walker and Yu Zhu (2006), *Who Benefits from Child Benefit?* Warwick Economic Research Papers, No. 749.

Camerer, Colin, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue and Matthew Rabin (2003), *Regulation for Conservatives: Behavioral Economics and the Case for 'Asymmetric Paternalism'.* University of Pennsylvania Law Review, Vol. 151, 1211-1254.

Card, David and Michael Ransom (2007), *Pension Plan Characteristics and Framing: Effects in Employee Savings Behavior.* NBER Working Paper no. 13275.

Carter, John and Michael Irons (1991), *Are Economists Different, and if so, Why?* Journal of Economic Perspectives, 5, 171-177.

Choi, James, David Laibson, Brigitte Madrian, and Andrew Metrick (2003), *"Optimal Defaults".* American Economic Review (Papers and Proceedings), vol. 93 (2), 180-185.

Cronquist, Henrik and Richard Thaler (2004), *Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience.* American Economic Association Papers and Proceedings, May 2004

DellaVigna, Stefano (2009), *Psychology and Economics: Evidence from the Field.* Journal of Economic Literature, 47(2), 315–72.

Downs, Julie, George Loewenstein, and Jessica Wisdom (2009), *Strategies for Promoting Healthier Food Choices.* American Economic Review, 99(2), 159–64.

Epley, Nicholas, Dennis Mak and Lorraine Chen Idson (2006), *Bonus or Rebate?; The Impact of Income Framing on Spending and Saving.* Journal of Behavioral Decision Making, Vol. 19, No. 3, 213–227

Friedman, Milton (1953), *Essays in Positive Economics*. University of Chicago Press.

Glaeser, Edward (2006), *Paternalism and Psychology*. The University of Chicago Law Review, vol. 73, 133-156.

Gupta, Sunil (1988), *Impact of Sales Promotions on When, What, and How Much to Buy.* Journal of Marketing Research, 25 (November), 342-355.

Hoynes, Hilary and Diane Schanzenbach (2009), *Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program.* American Economic Journal: Applied Economics, American Economic Association, vol. 1(4):109-39.

Jacoby, Jacob (2000), *Is it Rational to Assume Consumer Rationality? Some Consumer Psychological Perspectives on Rational Choice Theory*. NYU, Ctr for Law and Business Research Paper No. 00-09.

Kahneman, Daniel, and Amos Tversky (1979), *Prospect Theory: An Analysis of Decision under Risk.* Econometrica, XLVII (1979), 263-291.

Kooreman, Peter (2000), *The Labeling Effect of a Child Benefit System.* American Economic Review, vol. 90(3), June, 571-583.

Kooreman, Peter and Henriëtte Prast (2010), *What Does Behavioral Economics Mean for Policy? Challenges to Savings and Health Policies in the Netherlands*. Revised version (15 March 2010), forthcoming in De Economist

Krosnick, Jon (1991), *Response strategies for coping with the cognitive demands of attitude measures in surveys.* Applied Cognitive Psychology, 5, 213-236.

Laibson, David, Andrea Repetto and Jeremy Tobacman (2001), *A debt puzzle*. NBER Working Paper 7879.

Loewenstein, George and Drazen Prelec (1992), *Anomalies in Intertemporal Choice: Evidence and an Interpretation.* Quarterly Journal of Economics, 107(2), 573-597.

Lyssiotou, Panayiota (2009), *Are Child Benefits Fungible? Evidence from a Natural Policy Experiment.* University of Cyprus.

McClure, Samuel, Jian Li, Damon Tomlin, Kim Cypert , Latané Montague, P. Read Montague (2004), *Neural Correlates of Behavioral Preference for Culturally Familiar Drinks.* Neuron, Volume 44, Issue 2, 379–387.

Milkman, Katherine and John Beshears (2008), *Mental Accounting and Small Windfalls: Evidence from an Online Grocer*. Harvard NOM Working Paper No. 08-024.

O'Donoghue, Ted and Matthew Rabin (1999), *Doing It Now or Later.* American Economic Review, March 1999, 89-103

Samuelson, Paul (1937), *A note on measurement of utility*. The Review of Economic Studies, Vol. 4, No. 2, 155-161

Schmidt, Henk (2009, November 6), *Can scientific thinking be taught?* Speech at the Erasmus University, Rotterdam, The Netherlands. Read the online version at: http://www.eur.nl/english/eur/corporate_events/dies_natalis/2009/can_scientific_thinking_be_taught/

Simon, Herbert (1957), *A Behavioral Model of Rational Choice*, in *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: Wiley.

Sunstein, Cass and Richard H. Thaler (2003), *Libertarian Paternalism is Not an Oxymoron*. University of Chicago Law Review 70, 1159–1202

Thaler, Richard (1999), *Mental accounting matters*. Journal of Behavioral Decision Making, 12(3), 183-206.

Thaler, Richard, Cass Sunstein and John Balz (2010), *Choice Architecture*. University of Chicago.

Tversky, Amos and Daniel Kahneman (1981), *The framing of decisions and the psychology of choice.* Science New Series, Vol. 211, No. 4481, 453-458.

Appendix A1: Copy of the e-mail sent to potential subjects

Dear student,

For my masterthesis I have to conduct an economic experiment, for which I need participants. I would appreciate it if you could participate in my experiment.
In the attachment you can find the word-document with the instructions. Please read them carefully and make sure you understand them. During the experiment, you will be asked to make several decisions and answer some questions about yourself.
I would like to ask you if you can fill in your answers in the document and then send it back to me before the 1st of June.
The instructions should be clear, but in case you have any questions, you can always send me an e-mail and I will reply as soon as possible.

Please note that you are not allowed to use a calculator during the experiment. If you do use one, your answers will be worthless to me.

Among all participants I will raffle three prizes of 15, 10 and 5 euros as described in the instructions.

Thanks in advance for your participation!

Greetz,
Laurens.

Note: I also sent a Dutch version of this e-mail to Dutch-speaking students, with the same text in Dutch.

Appendix A2: Instructions of the experiment

Welcome to this experiment.

In this experiment you will make several decisions. Please read the following instructions carefully. At the end of these instructions you will be asked some example questions to test whether you understood all the instructions. You can start with the real experiment as soon as you answered these questions correctly.

This experiment is about decision making. By making these decisions, you can earn happiness points. We ask you to imagine that these happiness points are valuable to you. To encourage active participation in this experiment, the experimenter will randomly draw 3 participants out of all participants one week after the experiment has finished. Out of these 3 participants, the participant who earned the most points in this experiment will win 15 euros. The number 2 and 3 will earn 10 and 5 euros respectively. Since there will probably be about 50 people participating in this experiment, the chance that you are selected is very reasonable.

One important thing to mention is that the use of a **calculator is not allowed** during this experiment.

The experiment:

At the beginning of each round, you will be given a budget of a certain amount of tokens. You will have to make a decision on how to spend this budget. You can spend your tokens on two goods: drinks in a bar and study books. By buying these goods, you can earn "happiness" points. **Please note that you will not get any happiness points for unspent budget.** These tokens will be lost after each round and will not earn you anything. Therefore, we assume that you want to spend your entire budget. Each new round you will get a new budget, independent of your decisions in previous rounds. So each decision you make is independent from previous decisions.

The price of one drink will be 2 tokens and the price of one study book will be 4 tokens. This will remain the same throughout the whole experiment. So:

Total price of purchase = (number of drinks x price per drink) + (number of study books bought x price per study book)

An example: if you decide to buy 4 drinks and 6 study books, the total price of this purchase will be (4x2) + (6x4) = 32 tokens.

The value drinks and study books have for you are expressed in "happiness" points. How valuable a specific amount of drinks or study books is for you will be denoted in a table during the experiment. Below this table, you will see the price of one unit of each good. Below you see an example of such a table.

Number of units	Drinks	Study books
0	0	0
1	10	20
2	19	36
3	27	48
4	34	56
5	40	60

Price of one drink: 2 tokens; Price of one study book: 4 tokens

In this example, if you buy 4 drinks, you will earn 34 happiness points with that purchase. The same mechanism holds for study books: if you buy 3 study books, you will earn 48 happiness points with that purchase. So if you buy 4 drinks and 3 study books, you will earn 34+48=82 happiness points. Note again that you do not earn any points with unspent budget. Since these happiness points are valuable to you, we expect that you want to spend your entire budget.

To summarize:

Total number of happiness points earned = number of points earned for purchased drinks + number of points earned for purchased study books

Now you will answer three questions to see if you understand the experiment.

1. Consider the table as shown in the example. If you buy 4 drinks and 5 study books, how many happiness points will you earn in total?

2. The price of one drink is 2 tokens. If your budget is 40 tokens and you want to spend your entire budget on drinks in a bar, how many drinks can you buy?

3. The price of one drink is 2 tokens and the price of one study book is 4 tokens. You decide to buy 12 drinks. How many study books can you still buy if you want to spend your entire budget of 40 tokens?

The answers to these questions are mentioned and explained on page 7. Should you still have any questions, please contact the experimenter.

==========

Now we will start with the actual experiment, where you can earn real points. The first 3 decisions will be made based on table 1, which can be found on page 5 of the instructions of this experiment. In this table you can see how many happiness points you will earn with a specific decision.

Please stay within the boundaries of your budget. If you do not, your results will be useless and therefore you will not be able to win any money in the random draw I will do after conducting this experiment.

Decision 1)

You have a budget of 40 tokens. The price of one drink is 2 tokens and the price of one study book is 4 tokens. You may now decide how many drinks and study books you would like to buy.

I would like to buy …. drinks

I would like to buy …. study books

Decision 2)

Now suppose that you get a wage rise in your weekend-job. This gives you extra income. As a consequence, your budget rises from 40 tokens to 60 tokens. So you now have a budget of 60 tokens that you can spend in any way you like. This extra income is the only difference compared to decision 1. All prizes and point values as mentioned in table 1 remain the same.

Please write down below how many drinks and study books you would like to buy.

I would like to buy …. drinks

I would like to buy …. study books

Decision 3)

Now suppose that you did not get the wage rise, so your own income is 40 tokens again. However, now the government wants to encourage students to buy more study books. They decide to give a subsidy (voucher) of 20 tokens to all students. However, you can only spend these tokens on study books. (For Dutch participants: imagine that the government gives you a "Boekenbon" of 20 tokens) If the amount of tokens you spend on study books is less than 20 tokens, you cannot use these tokens to purchase extra drinks. To summarize, you have a budget of 40 tokens that you can spend in any way you like. In addition to your budget, you get a study books subsidy of 20 tokens. Because prices remain the same, this just means that the first 5 study books you buy are costless to you. Point values as mentioned in table 1 remain the same as well.

Please write down below how many drinks and study books you would like to buy.

I would like to buy …. drinks

I would like to buy …. study books

Decision 4)

Now consider table 2, which can be found on page 6. Imagine that you have a full time job at a retirement home. Your boss requires you to work 40 hours per week doing jobs around the house for only two clients: Bassie and Adrian. Your boss does not care how many hours you work for each client, as long as you work 40 hours per week. You start the week with a good mood (translated into 410 happiness points), but you dislike both clients and per job you will do for them, you will lose some happiness points, as mentioned in table 2 (so by each job you do you will get less happy). An example: doing 7 jobs for client Adrian will "cost" you 50 happiness points and doing 5 jobs for client Bassie "costs" you 64 happiness points. Doing one job for Adrian takes you 2 hours and doing one job for Bassie (a much more demanding client) will take you 4 hours. So in your workweek you can do at most 20 (=40/2) jobs for Adrian and 10 (=40/4) jobs for Bassie, if you spend all of your time with that one client.

Your goal is to maximize your happiness: so you want to maximize the amount of happiness points you have at the end of the week. To summarize:

Number of happiness points earned = 410 – costs of working for Adrian – costs of working for Bassie

Remember: in total **you have to work 40 hours**. So the following has to be true:

(2 x "amount of jobs done for Adrian") + (4 x "amount of jobs done for Bassie") = 40

Please write down below how many hours you would like to work for each client.

I would like to work …. hours for client Adrian

I would like to work …. hours for client Bassie

Now please answer some background questions:

Gender: Male / Female

Age: …..

Study: ……

On a scale of 1 to 5, how much do you use mathematics in your study?

Not so much 1 2 3 4 5 Very much

On a scale of 1 to 5, how would you rate your math skills?

Not so good 1 2 3 4 5 Very good

On a scale of 1 to 5, how often do you participate in economic experiments?

Rarely 1 2 3 4 5 Very often

In what year of study are you in?

1 2 3 4 5 or higher

If you would like to participate in the lottery, please write down your e-mail address below.

……

Thank you for your participation!
You can now send your answers to laurensverhoef@hotmail.com or l.d.verhoef@uvt.nl.

Table 1

Number of units	Drinks	Study books
0	0	0
1	24	54
2	39	85
3	52	110
4	63	133
5	74	154
6	84	173
7	94	191
8	103	209
9	112	225
10	120	241
11	129	256
12	137	272
13	145	286
14	152	300
15	160	314
16	167	X
17	174	X
18	182	X
19	189	X
20	195	X
21	202	X
22	209	X
23	215	X
24	222	X
25	228	X
26	235	X
27	241	X
28	247	X
29	253	X
30	260	X

Price of one drink: 2 tokens; Price of one study book: 4 tokens

The value of each number represents the happiness points you will **earn** if you purchase the corresponding number of that product. For example, buying 10 drinks earns you 120 happiness points and buying 14 study books earns you happiness 300 points.

Note that in decisions 1, 2 and 3 it is never possible to buy more than 15 study books.

Table 2

Number of jobs	Adrian	Bassie
0	0	0
1	6	12
2	13	25
3	21	37
4	28	51
5	35	64
6	43	78
7	50	92
8	58	106
9	66	122
10	75	137
11	83	X
12	92	X
13	101	X
14	111	X
15	121	X
16	132	X
17	143	X
18	156	X
19	171	X
20	195	X

One job for Adrian takes you 2 hours; One job for Bassie takes you 4 hours.

The value of each number represents the points you will **lose** if you do the corresponding number of jobs for the client mentioned above. For example, if you decide to do 16 jobs for Adrian you will lose 132 points and if you do 2 jobs for Bassie you will lose 25 points. So, if your decision is to do 16 jobs for Adrian and 2 jobs for Bassie, you will earn 410 – 132 – 25 = 253 points. In total you will work 16x2 + 2x4 = 40 hours, so your boss is satisfied as well.

Since one job for Bassie takes you 4 hours and you only have 40 hours at your disposal, it is not possible to do more than 10 jobs for Bassie.

Answers to example questions

1. 94 happiness points
Explanation: You will earn 34 happiness points for the 4 drinks you bought and you will earn 60 happiness points for the 5 study books you bought, so in total you will earn 34+60=94 happiness points.

2. 20 drinks
Explanation: Your budget is 40 tokens and the price of one drink is 2 tokens. If you want to spend all 40 tokens on drinks, you can buy 40/2=20 drinks.

3. 4 study books
Explanation: You have bought 12 drinks. This costs you 12x=24 tokens. This means that you have 16 tokens left to buy study books. The price of one study book is 4 tokens, so you can buy 4 study books with your remaining 16 tokens.

Appendix A3: Definition of the variables used

1 if economist: dummy variable indicating whether a subject is an economist

1 if female: dummy variable indicating whether a subject is female

Age: the age of the subject

Mathematical ability: indicates on a scale of 1 to 5 how highly a subject rates his/her mathematical ability (this variable is not used in the analysis, but only to create a dummy indicating whether a subject has high mathematical ability)

Experience in economic experiments: indicates on a scale of 1 to 5 how often a subject has participated in economic experiments (this variable is not used in the analysis, but only to create a dummy indicating whether a subject has experience in economic experiments)

Year: indicates the year of study the subject is in. Note that the maximum value subjects could indicate is "5 or higher" and that this observation will be treated as if it were "5"

1 if high mathematical ability: indicates whether a subject has high mathematical ability (=4 or 5 on the variable "mathematical ability")

1 if experience in economic experiments: indicates whether a subject has at least some experience with economic experiments (2 or higher on the variable "Experience in economic experiments")

Percentage of points scored: indicates how many percent of the amount of points the subject scored. This variable is constructed with the following formula:

Percentage of points scored = (Actual number of points scored – 1019)/ 257*100%.
Note that 1019 is the minimum amount of points a subject could earn in this experiment and 257 is the maximum amount of points a subject could have left after 1019 had been subtracted from the actual score.

Treatment-difference: The amount of units of the subsidized goods bought in treatment 3 minus the amount of units of the subsidized goods bought in treatment 2.

1 if e-mailaddress submitted: indicates whether a subject has submitted an e-mail address

Note that in the experiment subjects were also asked how much they use mathematics in their study. This variable has not been used in the analysis, due to the fact that the answers to this question were too subjective: students who did the same study gave answers varying from 2 to 5. As a consequence, this variable would not be objective enough to be used in the analysis. The variable "mathematical ability" served as a good and more objective substitute, especially when creating a dummy that indicates whether a subject has high or low mathematical ability.

Appendix A4: The Wilcoxon matched-pairs signed-ranks test

Just like the standard t-test, the Wilcoxon test compares differences between two matched samples, so the variable I used in this test is calculated by subtracting the MPC in treatment 2 from the MPC in treatment 3. The first step that needs to be taken is to eliminate all differences of zero in the dataset. These observations will not be used in the Wilcoxon matched-pairs signed-ranks test. Having done that, one has to rank the absolute differences in the MPC between the two treatments. If two or more observations have the same rank, then they will all get the average rank of all their ranks, so if for example the ranked observations 3, 4, 5 and 6 all have the same absolute value, they will all get the rank 4.5 (=(3+4+5+6)/4). Then, after ranking all differences, the sign of the difference will be taken into account. If a difference is negative, then the corresponding rank will be given a minus-sign and if the difference is positive, the corresponding rank will be given a plus-sign. After having done this, the test statistic W can be calculated. W is the minimum of two functions: W_+ and W_-, where W_+ is the sum of all positive ranks and W_- is the sum of all negative ranks. The test statistic W will be compared to the critical value as mentioned in the table below. Only when the test statistic W is smaller than this critical value the difference between the two MPC's is significant.

	One-tailed level of significance			
	0.05	0.025	0.01	0.005
	Two-tailed level of significance			
	0.10	0.05	0.02	0.01
n				
5	0	-	-	-
6	2	0	-	-
7	3	2	0	-
8	5	3	1	0
9	8	5	3	1
10	10	8	5	3
11	13	10	7	5
12	17	13	9	7
13	21	17	12	9
14	25	21	15	12
15	30	25	19	15
16	35	29	23	19
17	41	34	27	23
18	47	40	32	27
19	53	46	37	32
20	60	52	43	37
21	67	58	49	42
22	75	65	55	48
23	83	73	62	54
24	91	81	69	61
25	100	89	76	68

Table A4: Critical values for the Wilcoxon matched-pairs signed-ranks test

Appendix A5: Robustness tests

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Dependent variable	**Percentage of points scored**			**Treatment-difference**		
1 if economist	5.098***	11.229**	10.397**	-1.003**	-1.369**	-1.156**
	(1.851)	(5.064)	(4.641)	(0.407)	(0.525)	(0.483)
1 if high mathematical ability	4.646**	9.895**	9.710**	0.010	-0.127	-0.038
	(1.759)	(4.327)	(4.293)	(0.387)	(0.448)	(0.449)
Age-18	0.576	1.251	1.200	-0.113	-0.182	-0.164
	(0.365)	(0.947)	(0.93)4	(0.080)	(0.098)	(0.098)
1 if female	0.250	3.946	3.802	-0.287	-0.479	-0.367
	(0.981)	(4.939)	(4.980)	(0.436)	(0.512)	(0.521)
Year	-0.376	-2.276	-2.176	0.114	0.164	0.122
	(0.678)	(1.730)	(1.701)	(0.149)	(0.179)	(0.178)
Year * "1 if econ"						
1 if experience in economic experiments		-1.429			0.479	
		(4.463)			(0.462)	
1 if e-mail address submitted			0.497			0.184
			(5.169)			(0.540)
Constant	90.065***	81.837***	81.192***	1.324**	1.698**	1.662**
	(2.584)	(6.380)	(7.341)	(0.569)	(0.661)	(0.768)
N	45	52	52	45	52	52
Adjusted R-squared	0.28	0.10	0.10	0.04	0.04	0.02

Table A5: Robustness tests. Note that regressions (i)-(iii), (iv)-(vi) and (vii)-(ix) use the same explanatory variables as regressions (2), (3) and (5) in table 2, respectively. Significance at 1%, 5%, and 10% level is denoted by *, **, and *, respectively.**