Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR WIRTSCHAFTS- UND
SOZIALWISSENSCHAFTEN

# Sunk-Cost Fallacy with Partial Reversibility: An Experimental Investigation

Corina Haita

**WiSo-HH Working Paper Series**
Working Paper No. 09
November 2013

WiSo-HH Working Paper Series
Working Paper No. 09
November 2013

# Sunk-Cost Fallacy with Partial Reversibility: An Experimental Investigation

Corina Haita, Universität Hamburg, Germany

ISSN 2196-8128

Font used: „TheSans UHH“ / LucasFonts

Die Working Paper Series bieten Forscherinnen und Forschern, die an Projekten in Federführung oder mit der Beteiligung der Fakultät für Wirtschafts- und Sozialwissenschaften der Uni-versität Hamburg tätig sind, die Möglichkeit zur digitalen Publikation ihrer Forschungser-gebnisse. Die Reihe erscheint in unregelmäßiger Reihenfolge.

Jede Nummer erscheint in digitaler Version unter
https://www.wiso.uni-hamburg.de/de/forschung/working-paper-series/

**Kontakt:**
WiSo-Forschungslabor
Von-Melle-Park 5
20146 Hamburg
E-Mail: experiments@wiso.uni-hamburg.de
Web: http://www.wiso.uni-hamburg.de/forschung/forschungslabor/home/

# Sunk-Cost Fallacy with Partial Reversibility: An Experimental Investigation*

Corina Haita [†a]

[a] Department of Economics
University of Hamburg
Von-Melle-Park 5, 20146 Hamburg, Germany
e-mail: corina.haita@wiso.uni-hamburg.de

November 13, 2013

**Abstract**

This study contributes to the literature attempting to document the sunk-cost fallacy in laboratory. I study the bias in a simple experimental setting, void of the previously acknowledged psychological roots of the sunk-cost fallacy. Under this design (i) the subjects have the possibility to partially recoup the investment in the initial course of action, (ii) the alternative course of action is made explicit and obvious and (iii) the associated returns of each course of action are deterministic. The sunk-cost fallacy hypothesis is not confirmed on the sample as a whole. However, I find evidence of its manifestation on subsamples for which I make the conjecture of having a better comprehension of the experimental task. Nevertheless, the bias appears to be independent of the size of the investment in the initial course of action. After controlling for mistakes in decisions and the effort put in the experimental task, I find that higher cognitive ability subjects are more prone to the bias. Finally, I argue that the previously acknowledged psychological drivers of the sunk-cost fallacy are not needed for the bias to manifest itself. Instead, I put forward the *realization utility* as the most likely underlying reason behind the manifestation of the sunk-cost fallacy under the current design.

**Keywords:** cognitive ability, experiment, realization utility, sunk-cost fallacy
**JEL Clasification:** C91, D03, D11, M41

*The research partially to this article was sponsored by Central European University Foundation, Budapest (CEUBPF). The theses explained herein are representing the own ideas of the author, but not necessarily reflect the opinion of CEUBPF.

†I am very grateful to my PhD advisor, Prof. Botond Kőszegi, for his guidance in developing this paper and to Mirco Tonin for his very helpful advice and suggestions. I am indebted to Prof. Daniela di Cagno who kindly facilitated my access to the CESARE lab and to Paolo Crosetto for helping with subjects' recruitment. All errors are mine.

# 1 Introduction

Normative economic theory indicates that only marginal costs and benefits should matter for decision making; therefore, costs incurred in the past are irrelevant for future marginal payoffs. Nevertheless, actual human behavior often violates this theory and people tend to account for historical costs. Thaler (1980) labeled people's failure to ignore sunk costs as the *sunk-cost effect*, also called *sunk-cost fallacy* or *Concorde fallacy* after the famous airplane development project of the British and French governments (Arkes & Ayton 1999).[1] In common language, the fallacy of sunk cost is the irrational behavior of "throwing good money after bad.", i.e. once found on a course of action to which they committed an investment (e.g. time, money, effort), people continue to stay on that course of action and invest even more resources despite it being unprofitable.

As Thaler (1980) points out, gathering field evidence to test the sunk-cost fallacy hypothesis is often hindered by problems of self-selection. Hence, evidence of the sunk-cost fallacy has been thus far limited to hypothetical scenarios and field experiments, while efforts for documenting it in laboratory are still surprisingly scarce and provide mixed evidence (Ashraf et al. 2010). On the one hand, hypothetical questions lack saliency and the subjects are always asked to *imagine* various scenarios based on which they *state* their decisions. On the other hand, field experiments are most of the time contextual and use real commodities (Harrison & List 2004). This interferes with subjects' unobserved Bayesian priors and experience in relation to the particular experimental context. At the same time, it is not unreasonable to conceive that (consumptions) decisions in the field are rarely individual. Hence, rather than observing individual behavior, very often the experimenter observes a group behavior (e.g. family or couple), which is affected by the relative bargaining power in the group's decision making.

In this paper I design a lab experiment in which subjects fall into three different groups depending on the size of the cost they pay for entering an initial course of action, i.e. the sunk cost. This cost can be either zero or positive. If the cost is positive, then it can be either low or high. Once found on the initial course of action, the subjects are offered a possibility to *revert* from it, towards accomplishing a given experimental goal. Moreover, they are *explicitly* given the alternative course of action to the initial course of action. In addition, the returns offered by each course of action are non-stochastic and they are known by the subjects. Therefore, the decision environment in this experiment is *void of ambiguity and uncertainty*.[2]

[1] Throughout, I will use these terms interchangeably.

[2] For example, Tan & Yates (1995) find that the explicit specification of the expected future returns

The experimental parameters set the full adoption of the alternative course of action and the total abandonment of the initial one as the rational choice. The results show a surprisingly small adoption of the alternative course of action even for those subjects who enter the initial course of action in a costless manner. This motivates the restriction of the analysis to subsamples which I conjecture to have a good comprehension of the experimental task. Under this restriction, indeed, I find that there is a significantly higher adoption of the alternative course of action in the group of subjects who entered the initial course of action free of charge relative to those who incurred a cost. In addition, cognitive ability appears to increase the treatment effect, i.e. higher cognitive ability subjects are more prone to the sunk-cost fallacy.

The motivation of this paper is twofold. On the one hand, as already discussed, the scarce and mixed evidence of the sunk-cost fallacy in lab experiments leaves room for yet another experiment attempting to document the fallacy. On the other hand, this study aims at shedding some light on the reasons behind its manifestation. The literature has identified several psychological channels for explaining the sunk-cost bias. First, *cognitive dissonance* makes it hard for people to admit they made wrong decisions in the past. Hence, in order to rationalize their past decisions they resort to *ex-post* self-justification by investing even more resources into an unprofitable course of action. Second, the literature on *ambiguity aversion* has pointed to the sunk-cost fallacy as being one of the anomalies generated by the aversion to uncertainty (AlNajjar & Weinstein 2009). Third, Thaler (1980) used the prospect theory, specifically the *loss aversion*, to explanation why people fall pray to the sunk-cost fallacy. Instead, the current study attempts to identify the sunk-cost fallacy in an environment void of the above-mentioned psychological drivers. Hence, the main contribution is that of showing that the previously acknowledged psychological channels for the manifestation of the sunk-cost fallacy are not necessary for the bias to make itself visible. Moreover, I put forward the *realization utility* (Barberis & Xiong 2012) as the most plausible mechanism behind its manifestation. The realization utility hypothesis suggests that people feel a burst of pain when a loss is realized and, therefore, avoid or delay the realization of this loss.

The experimental manipulation consisted of three groups, which differ with respect to the sunk cost incurred: one control group and two treatments. Subjects in each group were asked to make decisions regarding two assets, which I shall label as $A$, the initial asset, and $B$, the alternative asset. Each group was randomized between subjects, within each experimental session. The initial endowment of the control group consisted of 40 units of asset $A$ and 900 Experimental Euros (EE). The treatment groups

decreases the sunk-cost bias.

were endowed only with cash and, instead, were offered to buy the 40 units of asset $A$ from the experimenter. Hence, they had to make a binary choice between buying the 40 units of asset $A$ and not buying any unit of asset $A$. The two treatment groups represent two levels of sunk-cost, low and high sunk-cost, respectively. A subject in the low sunk-cost condition faced an ask price of 100 EE, while one in the high sunk-cost condition was asked 200 EE for each unit of asset $A$. However, before making their purchase decision the subjects were informed that they will be able to sell the 40 units back to the experimenter at the end of the session, for a price of 300 EE. This was meant to create a strong incentive for investment since it resulted in a sure profit of 200 and 100 EE, respectively by selling them back to the experimenter. The cash endowments for the treatment groups were chosen such that, if the subjects decided to buy the 40 units of asset A, after the purchase they were left with the same amount of cash as the initial cash endowment of the control group. Therefore, after the investment stage, the financial position of the subjects was the same across all groups: 40 units of asset A and 900 EE. This overcomes income effects in decisions.

The subjects who decided not to buy the initial asset kept their initial cash endowment and made no further decisions during the main sunk-cost experiment. Therefore, they were excluded from the sample of interest. Despite the strong incentive for buying the initial asset, there was still a small percent of subjects who preferred to keep the initial cash endowment. At this point, the reader might be concerned that this procedure introduces a self-selection problem. However, as I will argue, this is very unlikely to affect the results of the data analysis. First, it should be noted that at the time of the purchase decision the subjects did not know about the upcoming stages of the experiment. Therefore, their decision to buy the 40 units of the initial asset had no particular motivation, except that they could make a sure profit. Second, the subjects who chose not to invest in the 40 units of asset $A$ were asked to wait in the lab until the end of the session. Therefore, the opportunity cost of time is also excluded from the explanation of why some subjects choose not to invest. However, the only plausible explanation for their decision to choose the status-quo is that they had a poor understanding of the experimental task or they preferred not to engage in a cognitive effort entailed by the continuation of the experiment.

After the endowment with the initial asset was completed, either free of charge or for a cost, subjects in all groups were informed that their task in the experiment was to collect exactly 50 units of asset $A$ and/or $B$, in any combination. Moreover, they were told that assets $A$ and $B$ have the same redemption value. For achieving the 50 units of asset $A$ and/or $B$, the subjects in all groups had one single opportunity to engage in trade with asset $A$. Thus, they could buy more units of asset $A$,[3] sell all or some of them up to the endowment of 40 units, or keep all of the 40 units. Following the subject's trading decision, any missing unit for achieving the 50 units was automatically filled with units of asset $B$. Each unit of asset $B$ had a cost which was also known to the subjects. However, this cost was below the trading price of asset $A$, such that it made it optimal for the subjects in any treatment to sell all the 40 units of asset $A$ and buy the required 50 units of asset $B$. It should be pointed out that the trading price, which was given and known to the subjects before they made their trading decisions, was lower than either of the initial purchase prices, i.e. the sunk costs of the treatment groups. This allows for part of the initial investment to always remain sunk and to disentangle the rational motive for re-sale from the pure speculative one.

Hence, any extra unit bought from asset $A$ is interpreted as escalation on the initial course of action, while any unit sold is a step towards deescalation. Therefore, the confirmation of the sunk-cost fallacy would have the two treatment groups re-sell more of the initial asset than the control group, while the low sunk-cost treatment would re-sell more than the high sunk-cost treatment. Thus, varying the cost of the initial asset across the two treatment groups allows to test whether the sunk-cost fallacy is related to the size of the investment or rather to the mere fact of making an investment.

The results show no sunk-cost effect on the sample as a whole. This is mainly due to the fact that a large number of subjects were status-quo biased in the control treatment. In fact, only 27% of the subjects in the control group recognized the optimal course of action, i.e. sold only 40 units of asset $A$, while in all treatments together only 18% of the subjects recognized it. Moreover, the average units of asset $B$ used by the control group was 22 units, which is significantly below the optimal number of 50. This raises the concern that many of the subjects had a poor understanding of the experimental task. Therefore, I perform post-hoc analysis on subsamples which I conjecture to have had a better comprehension of the experimental task. Indeed, this analysis reveals a clear sunk-cost effect. Precisely, there are significant differences between the control group and each of the sunk-cost treatments, but the difference is not significant between the two sunk-cost treatments. This result suggests a sunk-cost effect which is independent of the size of the investment and it supports the finding of Ashraf et al. (2010) that paying something results in more use than paying nothing. In sum, conditional on understanding the experimental task, I find confirmation of the sunk-cost fallacy, even in the obvious and non-stochastic decision environment of this experiment, void of the

[3] Note that, given the target of 50 units, the maximum number of units of asset $A$ one could buy was 10.

previously acknowledged psychological roots of the bias.

Further, the suspicion that many subjects did not have a full grasp of the experimental task suggested a more detailed analysis of the effect of the cognitive ability, on the use of the alternative asset. The score of the cognitive ability test, which was administered as the second part of the experimental session, was used as a proxy for the comprehension of the experimental task. Indeed, regression analysis indicates that the cognitive ability score is a significant explanatory variable for the use of the alternative asset in the control group. This suggests that the (lack of) cognitive ability predicts mistakes in decisions, which, under the design of this experiment it translated into status-quo bias rather than sunk-cost bias. Moreover, the analysis shows a statistically significant effect of the interaction between the treatment and the cognitive ability score. This points to the fact that high-cognitive ability subjects might to be more prone to the sunk-cost bias. However, this interpretation cannot be separated from the alternative interpretation that exactly because they understood the experimental task better, the high-cognitive ability subjects are more likely to exhibit the bias.

Notwithstanding the fact that the results of this study might not be robust to replications, mainly due to subjects' poor understanding of the experimental task, they do open a discussion on the effect of the cognitive ability on the manifestation of the sunk-cost fallacy. To the best of my knowledge this explanation was not explored or accounted for in the previous experimental studies of the sunk-cost fallacy. This may also explain why the sunk-cost fallacy was not confirmed in other laboratory experiments or it was even found in reverse (e.g. Friedman et al. (2007)). The remainder of the paper is organized as follows. The next section presents the existing literature on the sunk-cost fallacy. In Section 3 I describe the experimental design and the procedure employed in the paper. Section 4 presents the data analysis and the results. In Section 5 I discuss some potential applications derived from the current design and Section 6 concludes.

## 2 Existing Literature

Arkes & Blumer (1985) is, perhaps, the most prominent paper documenting the sunk-cost fallacy. Their field experiment was able to capture the difference in behavior among three groups of theater season tickets buyers, who were randomly chosen to pay different prices: full price and two levels of discounted prices. The experiment shows that those who paid the full price of the ticket visited the theater more often during the season than those who paid a discounted price. Further, using situational questionnaires, the paper ascertains that people with training in economics are not less prone to failing to ignore the sunk cost than those without economics training. Similarly, my experiment allows to draw conclusions regarding differences in the manifestation of the sunk-cost fallacy for the subjects with economics versus non-economics training.

Considered to be the second field experiment investigating the sunk-cost fallacy, Ashraf et al. (2010) employ a randomized control trial in Zambia to test whether higher prices induce more product use. Their experimental design is able to isolate the sunk-cost effect from the self-selection effect, but they find no evidence of the sunk-cost effect, at least in the domain of health products used in their study. Their experimental manipulation is inspired by the unexpected random discount in the offer price manipulated by Arkes & Blumer (1985). However, unlike Arkes & Blumer (1985) and similar to my design, they also include a treatment with zero transaction price. Using this treatment they test the hypothesis of paying a positive price versus paying zero price and they find a sunk-cost effect, although not statistically significant. Interestingly, Ashraf et al. (2010) find evidence of the sunk-cost effect in households' answers to hypothetical questions, which is, however, inconsistent with households' actual behavior. This result seems to undermine the reliability of the findings from previous studies based on hypothetical questions, and reinforces the need for more laboratory experimental work in order to discriminate among the mixed evidence.

Further, Roodhooft & Warlop (1999) use hypothetical scenario questions in the field. They found that hospital managers significantly under-engage in outsourcing of catering services when they are told to imagine that prior to the decision of outsourcing, the hospital had an in-house production of meals. This effect is even stronger when they are told that in the event of outsourcing, they will have to make caterer specific investment. My design bears some similarities with their hypothetical Scenario 1 (the control) and Scenario 3 (the sunk-cost condition) in that it can also be applied to the decision to vertically disintegrate (e.g. holding the initial asset is equivalent to in-house production, while buying the alternative asset represents investment in switching to outsourcing). However, my design differs from their hypothetical scenario in two ways. First, outsourcing and in-house production can be used in combination, thus allowing to measure various degrees of the sunk-cost fallacy. Second, unlike in their scenario, my experiment allows for partially recouping the initial investment in the in-house production. The latter element should alleviate the sunk-cost fallacy.

It appears that most of the experimental literature investigating the sunk-cost fallacy makes use of contexts and situation, particularly in field studies where real goods are used. For this reason, the results obtained in such studies are rather confined to the context, the particular commodity used or the population treated. Along this line, Tan & Yates (1995) showed that the decision to escalate on an initial course of action is sensitive to the context in which the problem is formulated. Again using hypothetical scenario questions, the authors show that students who had prior instructions in sunk-cost principles did ignore it when the context of the problem was similar to the textbook examples. However, they failed to do so when the decision reflected a real-life situation such as choosing between two resorts near Singapore. By contrast, my study attempts to examine the sunk-cost fallacy in a neutral environment, void of context, and thus void of *a priori* believes, preferences or learned norms.

I am aware of only three studies investigating the sunk-cost fallacy in laboratory. First, using lottery valuations as a measure of escalation of commitment, Phillips et al. (1991) show that when the sunk costs are made more transparent, they are more likely to be ignored. Nearly half of their subjects failed to ignore the sunk cost when this was not explicitly paid, but it was only a verbal commitment. However, only 19% of their subjects exhibited the bias when the sunk cost was made more salient through the physical act of paying the lottery ticket (the sunk cost in their experiment). In the same study they show that market forces can significantly alleviate the sunk-cost fallacy. Second, Friedman et al. (2007) devised a computer game to isolate factors which determine the sunk-cost fallacy. They asked subjects to use mouse clicks from a given budget of clicks in order to discover "treasures" on "islands" on which they arrived by paying a sunk cost, which ca be either low or high. Their data fail to find a significant difference in the number of clicks on the "cheap" versus "expensive" islands. Most recently, Robalo & Sayagy (2013) document the manifestation of the sunk-cost fallacy on the use of information in decisions under risk. Their experiment shows that subjects over-weight costly information relative to free information and shift their beliefs towards extremes, which is not consistent with Bayesian updating. Finally, the authors argue that the loss aversion is a suitable explanation for the observed behavior of their subjects.

Apart from documenting the bias *per se*, the literature has also identified the main psychological drivers for the manifestation of the sunk-cost fallacy. These drivers are important if one intents to educate against the bias, since it is the cause rather than the symptom that one needs to treat. First, several studies argue that cognitive dissonance (or self-justification) is one root-cause for the manifestation of the sunk-cost fallacy. The reason is that people do not like to admit they made bad decisions in the past. Their need to appear rational to themselves and to others determines them to continue the initial course of action, despite the slim chances of success, in order to justify their past decision. Supporting this view, Staw (1976) finds that people are more committed to a previously chosen alternative if they are made responsible for that decision at an earlier point in time, especially if this prior decision had negative consequences. Similarly, Bazerman et al. (1984) find that being responsible for the existence of a sunk cost increases the amount of resources allocated for the continuation of the project, both at the individual and group level. Further, Arkes & Blumer (1985) found *ex post* rationalization of past decisions, i.e. the presence of the sunk cost generated inflated optimism. In the same vein, Knox & Inkster (1968) find that horse-race betters are more optimistic about the chances of success of their favorite horse immediately after committing a bet on it than before they made the bet. However, my experiment is designed such that self-justification cannot be an explanation for the sunk cost fallacy. The main argument comes from the way the information is supplied to the subjects. Specifically, when they decide on incurring the sunk cost, the only information they have is that this cost produces a sure return at the end of the experiment. Moreover, this return is explicitly specified and known in advance by the subjects. Since there is no deceiving in the experiment, the decision to invest is not only *ex-ante* optimal, but also *ex-post*. Therefore, there is no reason for self-justification on the side of the subjects.

Second, in their theoretical study, AlNajjar & Weinstein (2009) argue that, in a dynamic setting, a decision maker with Ellsberg preferences (this type of preferences is consistent with ambiguity aversion) fails to ignore sunk costs. The only experimental endeavor, of which I am aware, to explicitly investigate the role of the ambiguity aversion in leading people to honoring sunk costs is that of van Dijk & Zeelenberg (2003). However, they manipulated ambiguity with respect to the size of the sunk cost rather than with respect to the returns. They find that when the size of the sunk cost is ambiguous (no probabilities associated), the sunk-cost fallacy is lower compared to the case in which this size is specified. However, if ambiguity is manipulated with respect to the returns it is expected to increase people's tendency to account for sunk costs. Indeed, Tan & Yates (1995) find that the simple mentioning of the expected returns (the elimination of the ambiguity) reduces the sunk-cost bias. While the authors do not label this finding as being the effect of the ambiguity aversion, they discuss how the inclusion of information about the expected returns competes with the inclusion importance of the sunk cost such that it decreases its effect importance. Since everything is stated in deterministic terms, the design of my study is void of any ambiguity feature, both with respect to the sunk cost and the returns. Therefore, ambiguity aversion cannot account for the manifestation of the sunk cost fallacy in this study.

Finally, using prospect theory (Kahneman & Tversky 1979), Thaler (1980) explains how the psychic accounting system leads individuals to account for sunk costs. Because of the convexity of the utility function in the domain of losses, the decrease in utility from a loss is lower than the increase in utility from an equal sized gain. Therefore, as Arkes & Blumer (1985) argue, once the decision-maker is found in the domain of losses, she will be willing to take further risks in the hope of an eventual gain, i.e. people are risk lovers in the domain of losses. The experimental design of this paper does not put the subjects in the domain of losses, since after the initial investment their financial position is, in fact, higher than before the investment. Hence, my design also rules out the loss aversion as a root-cause of the sunk-cost fallacy.

# 3 Experimental Design

## 3.1 Treatments and parameters

Consider the situation in which a decision maker is pursuing a course of action towards achieving a given goal, at the time when she receives new information. At this point she learns that (i) for achieving the goal an alternative course of action is also available and (ii) she has the possibility of reverting from the initial course of action and partially recouping its investment. Given that according to the future costs and benefits it is optimal for the decision maker to abandon the initial course of action and adopt the alternative course of action, this experiment aims at investigating how much abandonment and how much adoption will occur. Failure to abandon the initial course of action is interpreted as sunk-cost fallacy, which, in this experiment, can occur in various degrees such that the sunk-cost fallacy can manifest itself in a continuous manner.

Formally, let us assume that there are two types of assets in the economy, asset $A$ and asset $B$. The goal of the decision maker is to accumulate $Q$ units of assets $A$ and $B$ in any combination. Furthermore, each asset has the same end unitary value $p$ regardless of its type. Next, let us suppose that the decision maker has already invested in $A_0 < Q$ units of asset $A$ for a unitary price $p_0^A$. Therefore, at the time of receiving new information the cost of purchasing the initial endowment of asset $A$, the amount $p_0^A A_0$, is sunk. When new information arrives, the decision maker learns that she can trade (sell or buy) units of asset $A$ for a unit price $p_1^A$, and that she can buy units of asset $B$ for the unit price $p^B$, such that she can collect the $Q$ units. Let $A_1$ be the number of units of asset $A$ she decides to sell ($A_1 < 0$) or buy ($A_1 > 0$), i.e. how much to revert from the initial investment and how much to escalate on the initial investment, respectively. Hence, the problem of the decision maker is to choose $A_1$ and $B$ such that to maximize her payoff composed of the revenue from holding the $Q$ units of asset minus the cost of buying units of asset $B$, minus the cost (plus the revenue) from trading the holdings of asset $A$ and minus the sunk cost :

$$\begin{aligned} \max_{A_1,\, B} \Pi =& pQ - p_B B - p_1^A A_1 - p_0^A A_0 \\ &\text{such that} \\ Q =& A_0 + A_1 + B \\ A_1 \geq & - A_0 \text{ and } B \geq 0 \end{aligned} \tag{1}$$

Substituting $B$ from the constraint, it turns out that the rational solution is:

(i) if $p^B > p_1^A$, then $A_1 = Q - A_0$ and $B = 0$

(ii) if $p^B < p_1^A$, then $A_1 = -A_0$ and $B = Q$

Case (i) says the it is optimal for the decision maker to keep the initial asset and buy more units of asset type $A$ such that to complete the $Q$ units. However, since case (ii) predicts the abandonment of the initial investment, thus allowing to identify the sunk-cost bias, the parameters of the experiment are chosen accordingly. Hence, the cost of the alternative course of action, $p^B$, was chosen to be lower than the re-sale price, $p_1^A$ of the initial course of action. Moreover, in order for part of the initial investment to always remain sunk, this price must always be below the initial purchase price, i.e. $p_1^A < p_0^A$. The values of all experimental parameters are shown in Table 1 of Appendix A.

The experiment consists of three manipulations regarding the unit price, $p_0^A$, of the initial investment in $A$, as shown in Table 1. While all treatments received the same number of units of asset $A$,[4] the price paid for each unit was different. Subjects in treatments T100 and T200 were given an initial cash endowment and were asked to invest part of this endowment in acquiring 40 units of asset $A$. They paid 100 and 200 Experimental Euros (EE), respectively, for each unit of asset $A$. Moreover, the initial investment was a sizable amount from the initial cash endowment, i.e. 80% and 90% for T100 and T200, respectively. These are the sunk-cost treatments. Subjects in the control treatment T0 received the endowment of 40 units of asset $A$ free of charge. In order to avoid income effects, the initial cash endowments were such that, following the investment, subjects in all treatments had the same financial position: 40 units of asset $A$ and 900 EE. Thus, apart from the free endowment of 40 units of asset A, subjects in

[4]In order to avoid focal point effects or *a priori* preferences, assets' labels were randomized within treatments. Thus, in the same treatment, some subjects started with asset $A$ and had $B$ as the alternative asset, while other subjects started with asset $B$ and had asset $A$ as alternative. However, for the sake of exposition, I will continue using $A$ for the initial asset and $B$ for the alternative asset.

treatment T0 also received a cash of 900 EE. This amount of cash was chosen such that to allow for the purchase of the extra 10 units of asset $A$ for achieving the total of 50 units, in case the subject chose to fully escalate on the initial course of action.

Under the design of this experiment, the sunk-cost fallacy hypothesis can be formulated as follows:

**Hypothesis.** *Subjects in treatment* $T0$ *use more units of the alternative asset* $B$ *than subjects in treatment T100, who, in turn, use more units than those in treatment T200.*

Thus, the sunk-cost hypothesis is confirmed if the subjects in treatment T0 sell more units of the initial endowment than the subjects in treatment T100, who, in turn, will sell more than those in treatment T200.

### 3.2 Procedure

Seven experimental sessions were conducted during December 2012 and April 2013 in the CESARE laboratory at LUISS Guido Carli in Rome. A total of 153 subjects participated in the experiment and they were recruited online through the ORSEE system (Greiner 2004), from the subjects pool of the laboratory composed of students at LUISS. The participants belonged to Economics, Business Administration, Political Science, Communication Science and Law majors, out of which 67% were Economics or Business students. No subject participated in more than one session, i.e. the analysis of the treatment effect is carried out in a between-subject design. There were between 18 to 26 subjects in each session, with an average of 22 subjects per session. Each experimental session lasted for approximately one and a half hour, including subjects' payment. The interface of the experiment was programmed in z-Tree (Fischbacher 2007). Snapshots of the experimental screens, containing the instructions, are presented in Appendix B.

Upon arrival to the lab the subjects were randomly assigned to a working station. All subjects in the room saw the introductory screen in Figure 1. The instructions on this screen were read aloud by the experimenter. Subsequently the subjects followed the instructions on their respective screens and took their decisions individually. Every experimental session had three parts. The first part consisted of the main sunk-cost experiment. Subjects' payment for this part of the experiment was based on an exchange rate of 1500 EE for 1 euro. In the second part, all subjects (including those who chose not to invest in the main sunk cost experiment) answered the Holt & Laury (2002) (henceforth, HL) risk preference elicitation questions which were payment-incentivized (see Figure 8). One pair of lotteries was then randomly selected and the subjects were paid according to the lottery they chose in that pair. The exchange rate was 1000 EE for 1 euro. Finally, all subjects answered a cognitive quiz composed of five questions, with the value of 0.5 euro for each correct answer. The first three questions in the quiz were the cognitive reflection test (CRT) of Frederick (2005) and the final two questions were selected from the math quiz in Benjamin et al. (2006). The complete quiz can be found in Appendix C. Cumulative earnings from all three parts of the experiment ranged between 5.1 and 17.6 euro, with an average of 13.7 euro per subject. At the end of the experiment the subjects answered demographic questions and they had the chance to give reasons for their decisions in an open-answer question.

In the main sunk-cost experiment, the three treatments were randomized within sessions, with positive probability of each treatment, in each session. After the introductory screen, the subjects were presented with a screen in which they were informed about their initial endowments (see Figure 2). Thus, those in T0 were informed that they were endowed with 40 units of asset A and 900 EE, while those in T100 and T200 were endowed *only* with cash in the amount of 4900 and 8900 EE, respectively (see Table 1). Unlike the subjects in T0, those in T100 and T200 had an additional screen in which they were offered to invest in exactly 40 units of asset $A$ for a price of 100 and 200 EE respectively (see Figure 3). This was a take-it-or-leave it offer. If the subejcts chose to invest, they continued the experiment with further decisions. If they chose not to invest, they were asked to wait quietly in their seats until the end of the session. In order to make the investment salient, for those who invested the following screen emphasized the change in their cash account and the holding of the 40 units of asset $A$ (see Figure 4). In fact, during the whole experiment, the subjects could see their current financial position on the top-right corners of their screens.

Further, all subjects in T0 and those who invested in T100 and T200 faced the Trade decision (see Figures 5 and 6). At this stage they were told that they had to collect a total of 50 units of assets $A$ and/or $B$, in any combination, and that they had the opportunity to trade (sell or buy) units of asset $A$, or keep the units they already posses. Trading was possible for a price which was randomly drawn from the uniform interval 50 to 90 EE, before they made their decision. The instructions emphasized that they had only one opportunity to trade and that asset $B$ was automatically assigned given their trading decision such that in the end they would hold 50 units of assets $A$ and/or $B$. Finally, the subjects were informed that each unit of asset $B$ cost 30 EE and that the redemption value of each unit was 300 EE regardless of the type of asset, $A$ or $B$. The experiment lasted for one period only.

# 4 Results

The data analysis is based on the sample composed of all subjects participating in the free treatment and those subjects in the sunk-cost treatments (T100 and T200) who decided to invest in the initial asset. I do not believe that this procedure poses any problem of self-selection for two reasons. First, the design creates a strong incentive for investment, since it was obvious that investing brings a sure profit. Second, at the time of the decision to invest the subjects were not aware of how and for what they could use the asset they bought. All they knew was that they could re-sell the asset for a higher price than the price paid for acquiring it. Nevertheless, out of the 105 subjects in the two sunk-cost treatments, 11 subjects (or 10%) still chose to keep the initial money endowment. The non-investors were significantly more in the T200, 8 out of 56, compared to 3 out of 49 in T100, reflecting the higher cash offered in this treatment but also the smaller profit from investing. One would expect risk aversion to be responsible for their decision of not investing, to the extent that they did not want to engage in a game of which they did not have enough information. However, this does not seem to be the case. Within the sample of non-investors, out of those who made consistent choices in the HL lotteries, only one subject exhibited risk aversion, switching to the risky lottery only at the eighth pair. The majority of the subjects who did not invest were closer to risk neutrality, switching at the fourth or fifth pair. In fact, the sample of non-investors has on average a lower number of choices for the "safe" lottery than the sample who invested: 5.13 compared to 5.34. However, the only observable trait in which this sample seems to differ from that of those who invested, is the cognitive ability. Their average cognitive score was 2.27 correct answers as compared to 3.13 correct answers given by those who invested.[5] Hence, one possible explanation for the decision of not investing despite its obvious benefit, is the refusal to engage in a cognitive effort entailed by the continuation of the experiment. Thus, the decision to buy the asset is unrelated to the intention of using it. This should minimize concerns of self-selection.

Having established this, the final sample of analysis consists of 142 subjects distributed as 48, 46 and 48 subjects in T0, T100 and T200 respectively. The variable of interest for testing the hypothesis of this study is the number of units of asset $B$ used by the subjects in their task of gathering 50 units of asset $A$ and/or $B$. The possible values of this variable range from 0, indicating full escalation of commitment, to 50, meaning full abandonment of the initial course of action. Precisely, values from 0 to 10 indicate escalation of commitment (the subject's trading decision was to buy more units

[5] Recall that the total number of questions in the cognitive quiz was 5.

of the initial asset or passively keep them), while values from 11 to 50 indicate partial to full deescalation of commitment (the subject recognized the optimality of selling at least one unit of the initial asset). Hence, the set-up of this experiment allows for a continuous manifestation of the sunk-cost fallacy, in the sense that the subjects are not asked to fully abandon or fully escalate on the initial investment, but they can choose intermediary positions.

In the formal analysis I use the non-parametric Wilcoxon-Mann-Whitney (WMW) test and I report exact $p$-values. Corresponding to this test, the alternative hypothesis of my study is that the number of units used from the alternative asset by the T0 subjects is greater than that used by the T100 subjects, which, in turn, is greater than the number of units used by the subjects in the T200 condition. However, before proceeding to the main analysis, let us first notice whether there are any differences in subjects' individual characteristics across treatments. Descriptive statics by treatments are presented in Table 2 of Appendix A. There are no significant differences among the three treatments with respect to the individual characteristics, except for the gender difference between T0 and T200 (WMW $p$-value$= 0.0371$).

Figure 9 in Appendix D shows the kernel density estimates of the distribution of the number of units of asset $B$ used by each treatment. The first issue to note is that most of the subjects were unable to recognize the optimal decision. In fact, across all treatments, only 68% of the subjects used between 10 and 50 units of the alternative asset and only 18% recognized the optimal strategy of using 50 units. Next, it is readily visible that the subjects in T100 and T200 are more likely to have escalated on the initial commitment relative to those in the control treatment T0. Moving towards right, this order it reverted: the control treatment T0 has a higher likelihood of using units of the alternative asset closer to the optimal amount. However, there are only 27% of the subjects in the control treatment T0 who chose to use the maximum possible units of the alternative asset and only 11% and 17%, respectively in treatments T100 and T200. At the same time, only 15% of the subjects in T0 escalated on the initial course of action by using 0 units of the alternative asset compared to 24% and 19% in treatments T100 and T200, respectively. These results show that there are slight treatment differences: those subjects who did not pay for their endowment of the initial asset were more prone to abandon the initial course of action and follow the alternative course of action compared to the sunk-cost treatments, T100 and T200.

Table 3 in Appendix A shows the averages of the units used from the alternative asset, as well as the averages of units sold (abandonment of the initial course of action) and bought (escalation on the initial course of action), by treatment. Thus, column (1) is the mere consequence of columns (2) and (3), respectively. While it appears that the subjects in the control treatment were more willing, on average, to disregard the sunk cost and embrace the alternative course of action (22.44 units in the control as compared to 15.7 and 19.46 units in the two treatment groups respectively), non-parametric testing does not show statistically significant treatment differences. According to the 1-sided WMW test, there is a difference between T0 and T100 ($p$-value$= 0.0277$). However, this does not carry over for T0 versus T200 ($p$-value$= 0.2203$). Moreover, when testing for the joint hypothesis that $T0 \geq T100 \geq T200$, the Jonckheere trend test fails to reject that the three samples come from the same population (1-sided $p$-value$= 0.237$). In sum, the hypothesis of the sunk-cost effect cannot be confirmed on the sample as a whole.

However, as in Ashraf et al. (2010), I conduct a post-hoc test of whether there is an effect of a positive investment as opposed to zero investment, i.e. pooling T100 and T200 together and testing it against T0. Indeed, I find that, on average, the control treatment used the alternative asset more than the sunk-cost treatments together (22.44 units as opposed to 17.63 units) and this difference is statistically significant at 10% level (1-sided WMW test $p$-value$= 0.061$). This suggests that the sunk-cost fallacy might be a bias due to the mere fact of making an investment and not to the actual size of the investment. The effect seems to be somewhat stronger than the one found by Ashraf et al. (2010).

It should be noted that the weak result obtained on the sample as a whole might be due to several experimental implementation flaws and, therefore, it might not be robust to replications. First, the gender distribution on the three treatments is unbalanced, with the over-representation of males in the T200 treatment. Second, the English language proficiency of the subjects is questionable.[6] Third, there is indication of low opportunity cost of subjects' time and low effort dedicated to the experimental task. Particularly, when asked to explain the reasoning behind their trading decision in the questionnaire administered at the end of the experiment, many subjects admitted that they had no specific motivation or that they made the decision randomly.[7] Therefore, for the remaining of this section I conduct post-hoc tests on subsamples which I will argue to be less affected by frivolous decisions.

[6] Most of the subjects were Italian students and the ORSEE system for subjects recruitment did not give the experimenter the option to select students in English taught courses. However, the invitation to participate in the experiment was sent out in English and it explicitly stated that the understanding of the English language was a must.

[7] Among their answers I can quote: "I don't know", "It is a passion", "it was random", "curiosity", "for a new experience", "no reason".

## 4.1 Consistent HL lottery choices

One sensible way to account for frivolous decisions is to disregard from the analysis those subjects who made inconsistent choices[8] in the HL lottery menus for risk preference elicitation. Thus, in this subsection I use consistency in the HL lottery menus choices as a proxy for reliability of subjects' decisions in the main sunk-cost experimental. The rate of inconsistent choices in my sample is 27% (or 39 subjects out of 142), which is about twice as much as in the original HL study.[9] Charness & Viceisza (2011) find a very high percentage (about 75%) of inconsistency in the responses to the HL risk elicitation task among subjects in rural Senegal. The authors argue that this inconsistency is due to a low level of understanding of the task or frivolity in responses. Therefore, it seems reasonable to believe that those subjects who made consistent choices in the HL lottery menus, were also more likely to have made more conscious decisions in the main sunk-cost experiment. Those who made inconsistent choices, on the other hand, were more likely not to have taken the experimental task seriously due to a low opportunity cost of time or to have had a poor understanding of it.

The HL risk aversion questions were provided with real payment for one randomly selected pair of lotteries. Although this part of the experiment followed right after the main sunk-cost experiment, there is no evidence that the inconsistent answers were due to one treatment or another, i.e. no treatment driven "attrition". Indeed, the $p$-values of the WMW test of differences between T0 and T100, between T0 and T200, and between T100 and T200 regarding the inconsistent answers are 0.4964, 1 and 0.4964, respectively, indicating that there are no distributional differences among the treatments with respect to the consistency of the lottery choices. Therefore, I can assume that the subjects with inconsistent HL choices are randomly distributed across the three experimental treatments and that the most likely explanation for their inconsistency is the low level of attention and/or effort exercised in the experimental task.

Figure 10 in Appendix D shows the kernel distribution of the units used from the alternative asset, both for the subjects with consistent and those with inconsistent HL choices. While the inconsistent subjects do not exhibit a clear pattern, the subsample of consistent subjects shows treatment differences between the control and each of the treatment groups. Hence, it is worth looking at the behavior of the subsample with consistent HL lottery choices alone. Table 4 in Appendix A shows the treatment averages

[8] An inconsistent choice means switching from lottery A (the "safe" lottery) to B (the "risky" lottery) and back to A, or backwards choices switching from the "risky" to the "safe" lottery, or choosing lottery A in the last row of the menu. I classify those who chose only lottery B as consistent.

[9] Other studies show inconsistency between 10 and 15 percent (Charness & Viceisza 2011).

for this subsample as well as the number of subjects in each treatment. As it can be seen, the treatment are relatively balances concerning the sample sizes and the treatment differences are sharper than on the sample as a whole. Precisely, the 1-sided WMW test shows that the number of units used of the alternative asset in the control group T0 is statistically larger, at conventional levels, than both those used by T100 and T200 groups ($p$-value=0.018 and $p$-value=0.056, respectively). The same test suggests no significant difference between T100 and T200 ($p$-value=0.481). This is in line with the findings based on hypothetical questions of van Dijk & Zeelenberg (2003), who found no statistical difference between their high and low sunk-cost groups. However, a test of the joint hypothesis that $T0 \geq T100 \geq T200$, rejects the null that the use of the alternative asset was the same across the three treatments (Jonckheere's trend test 1-sided $p$-value= 0.064).

Despite the fact that the size of the sunk cost does not appear to matter for the manifestation of the fallacy, the statistical difference between the control and the two treatment groups in the direction of the sunk-cost fallacy is surprising. One would expect that subjects who exhibit consistent preferences, in this case with respect to their risk attitudes, are also more likely to make rational decisions in other domains (Choi et al. 2011). Therefore, this post-hoc sub-sample selection should, in fact, invalidate the sunk-cost fallacy hypothesis of this study, which does not seem to be the case. Hence, taken together, these results indicate a sunk-cost effect which is independent of the size of the sunk cost, which leaves the hypothesis of this study only partially confirmed.

### 4.2 Cognitive ability

Similar to the previous subsection and making use of the cognitive score test administered in the third part of the experiment, I conjecture that subjects with higher cognitive ability had a better understanding of the experimental task than those with low cognitive ability. Recall that the cognitive test involved real payment for each correct answer. It consists of five questions with the first three questions representing the cognitive reflection test (CRT) of Frederick (2005) while the final two are mathematical reasoning questions extracted from the cognitive quiz in Benjamin et al. (2006).[10]

Hence, I split the sample in "low" and "high" cognitive ability according to the cognitive score of the 5-question cognitive quiz. I qualify the lowest 75th percentile (3 correct answers or less) as low cognitive, and the highest 25th percentile (4 or 5 correct answers) as high cognitive.[11] Figure 11 in Appendix D shows the density functions

[10] See Appendix C for the complete test.

[11] Note that according to this manner of splitting the sample, a high-cognitive subject is one who

of the units used from the alternative asset by treatment, separately for each of the two subsamples. Indeed, within the low-cognitive subsample there are no perceivable differences among the three treatments. Moreover, the average number of units used from the alternative asset is significantly below the optimal level in all treatments, including the control (see Table 5 in Appendix A for treatment averages and subjects' distribution across treatments). In fact, within the control group, there is a highly statistically significant difference between the low cognitive and the high cognitive group (WMW test $p$-value= 0.000). This reinforces the assumption that the low-cognitive group had a poor understanding of the experimental task. Therefore, this sub-sample does not appear suitable for the analysis of the treatment effect.

Focusing on the high-cognitive sub-sample it is easy to see that there are a few differences among the treatments. First, the distribution of the control group is shifted to the right, showing that, on average, this group made more use of the alternative asset than the two treated groups. Moreover, this distribution starts at 10 units, which indicates that subjects in this group recognized the optimality of switching to the alternative asset, even if they did not do it entirely. This does not hold for T100 and T200. Second, the share of the subjects who switched completely to the alternative asset is significantly higher in the control group than in any of the two treatments. These observations are confirmed by the WMW test. The $p$-values are provided in Table 6. They show a clear treatment effect of paying for the initial course of action as compared to the two treatment groups. However, as in the case of the analysis based on the sample of consistent HL answers, there is no statistical difference between the two sunk-cost treatments.

The results of this exercise also alleviate concerns of selection bias resulted from the existence of a group of subjects who chose not to invest. According to their cognitive score, these subjects would have fallen into the low-cognitive ability group had they invested. Therefore, they do not affect the results found for the high-cognitive subsample. Moreover, given their cognitive ability score, they are more likely to have had a poor understanding of the experimental task. Consequently, the extent to which the selection problem effects the overall results of the experiment is that more subjects would have been status-quo biased and, thus, inconclusive for the analysis.

Summarizing, conditional on subjects' making conscious and non-frivolous decisions, the sunk-cost hypothesis of this study is partially confirmed. In particular, conditional on high cognitive level, there is a sunk-cost bias, which is, however, independent of the actual size of the initial investment. In Section 4.4 I investigate in more detail the role of the cognitive ability, using regression analysis.

answered correctly at least two questions of the CRT.

## 4.3 Economists versus non-Economists

Because the literature has discussed the difference between the behavior of Economics versus non-Economics, or Accounting versus non-Accounting, students (Tan & Yates 1995, Arkes & Blumer 1985), I further split the sample according to the major of studies. In the "Econ" sample I include the subjects majoring in Economics or Business Administration and in the "non-Econ" sample I include all the other subjects. For each of these subsamples, Figure 12 in Appendix D shows the kernel distribution of the unis used from the alternative asset. As this figure shows, for the subsample of Economists ($N$=92, distributed as 32, 26 and 34 on T0, T100 and T200, respectively) the treatment differences appear to be more consistent than on the sample as a whole. Indeed, the 1-sided WMW test shows significant statistical differences, suggesting that the subjects in the control condition T0 made use of the alternative asset more than both those in the T100 treatment ($p$-value= 0.045) and those in T200 treatment ($p$-value= 0.022), respectively. Consistent with the findings from the previous subsections, the test does not detect any difference between the two treatment groups, T100 and T200 (2-sided $p$-value= 0.860), reinforcing the idea that the sunk cost fallacy might be independent of the size of the sunk cost. Nevertheless, the multiple hypothesis testing using the Jonckheere trend test shows a significant trend in the use of the alternative asset across the treatments (1-sided $p$-value= 0.024).

Further, on the sample of Economists, I perform the test of paying nothing versus paying something, i.e. testing for the difference between the free condition (T0) and the sunk-cost condition (pooling T100 and T200 together), and I find a significant effect of a positive sunk cost ($p$-value= 0.015). This result suggests again, as in the case of the subsample of consistent HL choices, that the manifestation of the sunk-cost fallacy is independent of the size of the initial investment, at least on the subsample of Economists.

The subsample of non-Economists ($N = 50$), on the other hand, is distributed as 16, 20 and 14 subjects on the three treatments, T0, T100 and T200, respectively. Interestingly, this subsample shows a reverse sunk-cost effect. The T200 treatment has significantly higher values than the control (WMW $p$-value= 0.036) and the T100 treatment (WMW $p$-value= 0.004), respectively, but the control and T100 treatment are not different from each other with respect to the number of units used from the alternative asset (2-sided WMW $p$-value= 0.3355). The surprising result among the subsample of non-Economists may be due to the unbalanced cognitive abilities across treatments within this subsample. The subjects in T200 seem to have significantly higher cognitive abilities than both those in T0 (WMW $p$-value= 0.025) and those in T100 (WMW

$p$-value= 0.1095). No such differences exist on the subsample of Economists.

Finally, it should be noted that the "Econ" subsample has significantly higher cognitive abilities than the "non-Econ" subsample (WMW $p$-value= 0.013). Therefore, consistent with the conjecture from the previous subsection, it seems plausible to believe that the non-Economists might have been more confused by the experimental task, especially because it was formulated in the language of an economic problem, involving assets and prices. From this reason, the results obtained on the subsample of non-Economists should be regarded with caution. Instead, more confidence should be put on the results obtained on the subsample of Economists with respect to verifying the sunk cost fallacy hypothesis.

### 4.4 The role of cognitive ability

The previous non-parametric analysis pointed to the fact that cognitive ability may play a role in the manifestation of the sunk-cost fallacy. Therefore, in this subsection, I investigate this possibility, resorting to regression analysis. Cognitive ability has been found to be responsible for many behavioral biases such as the conjunction fallacy, anchoring, base rate fallacy, conservatism and overconfidence (Oechssler et al. 2009, Hoppe & Kusterer. 2011), but also for the risk aversion and impatience (Frederick 2005, Benjamin et al. 2006, Dohmen et al. 2010). However, to the best of my knowledge, there is no experimental evidence about the role of the cognitive ability on the manifestation of the sunk-cost bias.

Table 7 presents the regression results of the effect of the cognitive ability on the number of units used from the alternative asset. Specification in column (1) shows the conditional treatment effects after controlling for the cognitive ability, proxied by the normalized cognitive quiz score. As the non-parametric analysis from subsection 4.2 showed, cognitive ability is responsible for mistakes in decisions. Since in this experiment mistakes can only induce less-than-optimal usage of the alternative asset, they can be confounded with the sunk-cost fallacy. Hence, after correcting for mistakes in decisions, the control group T0 appears to have used on average about 31 units of the alternative asset, which is still well below the optimal level of 50 units. Importantly, one standard deviation from the mean in the cognitive quiz score, reduces the mistake by 5 units and this is statistically significant (see the coefficient on "Cognitive" variable in Table 7). This suggests that people with higher cognitive ability are more likely to recognize the optimality of the alternative course of action, or more unlikely to make mistakes in decisions. The coefficients on T100 and T200 in column (1) show the treatment effects after controlling for cognitive ability. The negative signs of these coefficients show that there are, indeed, treatment effects in the expected direction. However, the coefficient is only marginally significant for T100 and insignificant for T200.

All these coefficients are robust to controlling also for the effort put in the experimental task, proxied by the time the subjects used to make their decisions in the "Trade" stage (see column (2)). This is a z-Tree recorded time and it includes the time used for reading the instructions in the "Trade" stage, cumulated with the time used for making their decisions at this stage. Although the effort variable turns out not to be statistically significant, it shows that putting more effort in making the decision leads to a greater use of the alternative asset, up to a turning point after which "over-thinking" leads to suboptimal decisions.

Further, in column (3) I let the treatment effects vary with the cognitive ability. While the conditional treatment effects do not change significantly, the interaction terms of the treatment dummies with the cognitive quiz score show that the treatment effect increases in the cognitive ability. However, this result is driven by the fact that, as it was seen in the non-parametric analysis, the low cognitive group did not exhibit any treatment differences, while being generally status-quo biased.

In equations (1) to (3) I considered the complete cognitive quiz, including both the CRT and the mathematics questions. In the specification from column (4) I consider only the CRT questions as a proxy for the cognitive ability. While most of the coefficients have the same significance and magnitude as in equation (3), the effect of the CRT score on the units used from the alternative asset is stronger than that of the complete cognitive quiz score.[12] This suggests that innate cognitive ability, which is better captured by the CRT test, is more important than learned cognitive ability which is quantified by the mathematics questions.

Moreover, the coefficients of the interaction of the CRT score with the treatment dummies, are larger in magnitude than the corresponding coefficients in column (3). In addition, unlike in the specification including the score of the whole cognitive quiz, the effect of the cognitive ability on the sunk-cost bias of T200 turns out to be statistically significant, even if only marginally. Although the magnitude of the interaction coefficient with T100 is larger, a test of equality between the two coefficients fails to reject that the coefficients on the interaction terms of the treatment dummies and the CRT score are equal. The interaction terms show that, within the T100 treatment, one standard deviation above from the mean in the CRT score is equivalent to the use of additional

[12] Including both score of cognitive ability, the coefficient on the mathematics questions was not found significant.

1.2 units from the alternative asset. Similarly, within the T200 treatment, one standard deviation upwards from the mean of the CRT score induces the use of additional 4.6 units of the alternative asset.

In sum, conditional on the cognitive ability, there is a treatment effect, which is higher in magnitude and significant for the T100 treatment than for the T200 treatment. Moreover, the cognitive ability was found to be both economically and statistically significant for predicting mistakes in decisions. At the same time, higher cognitive ability subjects seem to be more prone to the sunk-cost effect. While this result seems to be the counterpart of the findings in the literature according to which infants and animals do not exhibit the sunk cost bias (Arkes & Ayton 1999), it must be regarded with caution. Due to the large number of subjects who seem to have had difficulties with understanding the experimental task, two competing explanations account for this result. First, it might, indeed, be the case that high-cognitive people misused the "don't waste" rule (Friedman et al. 2007) or, exactly because they understood the task they were more likely to exhibit the bias. The latter explanation is particularly appealing since mistakes due to the lack of understanding of the experimental task can only result in sub-optimal use of the alternative asset, under the design of this experiment.

### 4.5 Discussion

Despite the difficulty with subjects' understanding of the experimental task, the results of this experiment showed indication of the manifestation of the sunk-cost fallacy. Moreover, the bias made itself visible even under the experimental design sterile of the interference of its psychological drivers previously acknowledged by the literature. Hence, my study showed that ambiguity, cognitive dissonance and loss aversion are not necessary for the subjects to exhibit sunk-cost fallacy. Instead, the most likely explanation for the manifestation of the sunk-cost fallacy under the current experimental design can be adapted from the *realization utility* theory developed by Barberis & Xiong (2012). According to the authors, people feel a burst of pleasure when a gain is realized and a burst of pain when a loss is realized. In other words, people do not obtain utility only from consumption of goods and services, but also from the mere act of . This theory was confirmed using neural data of subjects' brain activity at the moment of submitting their trading decisions in an experimental stock market (Frydman et al. 2012).

Realization utility theory has been proven suitable for explaining the disposition effect in investors trading behavior, i.e. the greater propensity to sell a stock which increased in value relative to the purchase price and hold on to those which decreased in value relative to the purchase price. This is because people do not think of their investment history in terms of returns over the whole portfolio, but rather as separate investment episodes characterized by the name of the asset, the purchase price and the re-sale price (Barberis & Xiong 2012). This comes close to explaining why subjects in my experiment were reluctant to part with the initial asset when offered a price below their purchase price. Apparently, from the desire of avoiding the pain from the direct act of selling the initial asset for a loss (relative to the purchase price), the subjects were trapped into an unprofitable course of action, i.e. trapped into the sunk-cost fallacy. This theory is also in line with the results of an earlier experiment by Staw (1976), who found that people would invest more in a course of action with negative consequences (holding on an asset which decreased in value) than in one for which their prior decisions proved successful (parting with an asset which increased in value).

The partial reversibility nature of the investment is the element of the experimental design that makes realization utility the most pertinent explanation for the manifestation of the sunk-cost fallacy. While the experimental literature has not emphasized this element thus far, it is, nevertheless a realistic possibility. Below I discuss two applications in which the investment in the initial course of action can be partially recouped and I illustrate the manifestation of the sunk cost fallacy in each situation.

# 5 Applications

The experimental design of the current study has several applications. First, the design applies straightforwardly to a practical problem related to carbon emissions trading schemes (ETS). In such schemes, the environmental agency distributes a number of permits to the regulated firms. Most commonly, the allocation can be gratis, following a benchmarking rule (also called grandfathering allocation), or through an auction. Hence, in the former case, the regulated firms receive the permits for free, while in the latter case they have to pay. After the allocation is completed, firms can trade the permits among themselves in a secondary market. The experimental design of this paper captures the differences in trading behavior when carbon emissions permits are distributed for free as opposed to auctioning. Let us imagine that the polluter (the manager or trader of an energy company) invests in buying emissions permits when these are distributed in an auction. In terms of the design of this experiment this is equivalent to investing in the initial asset. Let us further suppose that a market for these permits opens and that also some emissions reductions technology becomes available. In the language of the experiment, this is to say that the initial asset can be sold and the alternative asset becomes available. Assuming the situation from the experiment, where the alternative asset (the emissions reduction technology) is cheaper than the re-sale price of the initial asset (the emissions permits), sunk-cost fallacy would result in sub-optimal adoption of the emissions reduction technology. By contrast, free allocation would provide a closer to optimum adoption of the emissions reduction technology.

Since different methods of initial allocation, particularly free allocation versus auctioning, can lead to different actual initial allocations, it turns out that even with frictionless markets, the distribution of property rights matters. Therefore, contrary to Montgomery (1972), the efficient equilibrium will not be achieved if managers are biased towards honoring previous investment in emissions allowances instead of recognizing the possibly cheaper emission reduction options or fuel switches. This will undermine the overall goal of an emissions trading scheme that is to spur green technologies. The results of this paper suggest that such concerns might not be undue.

Second, consider the situation of the decision to vertically integrate. The following case is close to the design of the experiment in this paper. Assume a vertically disintegrated company which has a contractual relationship with a supplier of an intermediate good. The supplier delivers the stock of the intermediate good at time $t$, according to the contract. At time $t + 1$, the head of the R&D department informs the manager of the company that the intermediate good can now be produced in-house and that the company already has the necessary technology, i.e. no additional investment is needed. Moreover, the in-house production can be done at a lower unit cost than the contracted price.[13] However, the contractual relationship cannot be broken immediately, such that the supplier will continue to deliver the intermediate good until $t + 2$. Hence, the price contracted for the delivery is the sunk cost. Nevertheless, a market for the intermediate good exists, such that the company could sell the current and future stocks of the intermediate good until $t + 2$. Importantly, the market price is above the in-house production cost, but below the contracted price.[14] Hence, apart from the short-run needs until the in-house production is set-up, the optimal decision of the manager would be to sell the intermediate goods supplied through the contract rather than using it in the production. Instead, the final output could be produced using the intermediate good from the internal production. If, however, the manager fails to recognize this alternative and delays the in-house production until $t + 2$, when the contractual relationship with

[13]This could also be regarded as a transfer price from a producer which is part of the same business group.

[14]This last assumption may seem unrealistic, but could be somewhat justified by the fact that the outsourcing contracts are signed for longer periods during which the market conditions could change, i.e. forward contracts which are common, for example, in the supply of electricity.

the supplier can be broken, then she had fallen into the trap of the sunk-cost fallacy.

## 6 Conclusions

The experiment of this paper documented the sunk-cost fallacy in a laboratory setting. Despite the simplicity of the design, the results showed that subjects had difficulties in finding the optimal course of action. The control group, in which no sunk cost was incurred, used on average about 30 units of the alternative asset, which is still well below the optimal level of 50 units.

While this study fails to confirm a sunk-cost bias on the entire sample of subjects, it does find evidence of the sunk-cost effect on subsamples which I argue to have a better comprehension of the experimental task, i.e. the subsample of subjects with consistent risk preferences, with high cognitive ability or Economics and Business majors. Thus, provided that the experimental task is well understood by the subjects, this experiment found manifestation of the sunk-cost fallacy, despite the obviousness of the alternative course of action, the deterministic decision-making environment and the partial reversibility of the initial investment, which characterize the design of this experiment. However, the sunk-cost fallacy was found to be independent of the size of the sunk cost. This confirms the findings of other studies that paying something results in more use that paying nothing. Finally, because the previously acknowledged psychological factors responsible for the sunk-cost bias are missing from the design of my experiment, it turns out that they are not needed for the sunk-cost fallacy to make itself visible. Instead, due the nature of the design of my experiment, I put forth the realization utility as the most likely psychological phenomenon responsible for the bias.

Although the results of this study might not be robust to replications, particularly due to subjects' poor understanding of the experimental task, they open the question of the effect of the cognitive ability on the manifestation of the sunk-cost fallacy. The regression analysis conducted in this paper showed that after controlling for the cognitive ability and the effort put in the experimental task, cognitive ability continued to have the effect of increasing the sunk-cost bias, i.e. the difference between the units used by the control groups and each of the treatment groups. In other words, the sunk cost bias increases in the cognitive ability. While this result seems counter-intuitive, it appears to be the counterpart of the evidence that animals and infants are not prone to the sunk cost-fallacy. However, there is an alternative interpretation to this results. In particular, the high cognitive ability subjects are more likely to understand the experimental task and, for this reason, they are also more likely to exhibit the bias. While, at this point, the author cannot distinguish between the two competing interpretation to the result that the treatment effect increases in the cognitive ability, more research effort it worth putting into understanding the relationship between the sunk-cost bias and cognitive ability.

Hence, several extensions and refinements of the design merit consideration for further research. First, in order to avoid noisy decisions, the manifestation of the sunk cost could be made discontinuous. In this situation, subjects would be given the option to choose between buying 10 units of the initial asset or selling the 40 units of the initial asset. Second, in the sunk-cost treatments T100 and T200, after the investment stage is completed, the subjects could be asked the unit price for which they bough the asset. This would have the effect of both making the investment more salient and checking whether the subjects understood the instructions. For the same purpose, in a final questionnaire, the subjects would be asked to remember the price for which they had bought the initial asset. This could allow the experimenter to verify whether the memory of the sunk cost is part of the information set of the subjects at the time of the trading decision. In the absence of this memory it would be hard to argue that decisions which appear to take into account the sunk cost are a consequence of a fallacy rather than a simple decision error. Finally, one should allow for learning. Allowing for multiple periods would give the subjects the opportunity to learn disregarding the sunk cost in decision-making. This is in itself an important research question, since real-life problems give people the opportunity to learn the optimality of their decisions by repeatedly facing similar situations and observing the outcome of their decisions.

# References

AlNajjar, N. I. & Weinstein, J. (2009), 'The ambiguity aversion literature: A critical assesment', *Economics and Philosophy* **25**(Special Issue 03), 249–284.

Arkes, H. R. & Ayton, P. (1999), 'The Sunk Cost and Concorde Effects: Are Humans Less Rational Than Lower Animals?', *Psychological Bulletin* **125**(5), 591–600.

Arkes, H. R. & Blumer, C. (1985), 'The psychology of sunk cost', *Organizational Behavior and Human Decision Processes* (35), 124–140.

Ashraf, N., Berry, J. & Shapiro, J. M. (2010), 'Can higher prices stimulate product use? Evidence from a field experiment in Zambia', *American Economic Review* **100**, 2383–2413.

Barberis, N. & Xiong, W. (2012), 'Realization utility', *Journal of Financial Economics* **104**, 251–271.

Bazerman, M. H., Giuliano, T. & Appelman, A. (1984), 'Escalation of commitment in individual and group decision making', *Organizational Behavior and Human Performance* **33**(2), 141–152.

Benjamin, D. J., Brown, S. A. & Shapiro, J. M. (2006), 'Who is ehavioralCognitive ability and anomalous preferences', *Working paper, Harvard University* .

Charness, G. & Viceisza, A. (2011), 'Comprehension and Risk Elicitation in the Field: Evidence from Rural Senegal', *IFPRI Discussion Paper 01135, International Food Policy Research Institute (IFPRI)* .

Choi, S., Kariv, S., Müller, W. & Silverman, D. (2011), 'WHO IS (MORE) RATIONAL?', *Working Paper 16791, NBER WORKING PAPER SERIES* .

Dohmen, T., Falk, A., Huffman, D. & Sunde, U. (2010), 'Are Risk Aversion and Impatience Related to Cognitive Ability?', *American Economic Review* **100**, 1238–1260.

Fischbacher, U. (2007), 'z-Tree: Zurich Toolbox for Ready-made Economic Experiments', *Experimental Economics* **10**(2), 171–178.

Frederick, S. (2005), 'Cognitive Reflection and Decision Making', *Journal of Economic Perspectives* **19**(4), 25–42.

Friedman, D., Pommerenke, K., Lukose, R., Milam, G. & Huberman, B. A. (2007), 'Searching for the sunk cost fallacy', *Experimental Economics* **10**(1), 79–104.

Frydman, C., Barberis, N., Camerer, C., Bossaerts, P. & Rangel, A. (2012), 'Using neural data to test a theory of investor behavior: An application to realization utility', *Yale ICF Working Paper No. 12-30* .

Greiner, B. (2004), 'The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics', *Working Paper Series in Economics 10, University of Cologne, Department of Economics* .

Harrison, G. W. & List, J. A. (2004), 'Field Experiments', *Journal of Economic Literature* **XLII**, 1009–1055.

Holt, C. A. & Laury, S. K. (2002), 'Risk Aversion and Incentive Effects', *The American Economic Review* **92**(5), 1644–1655.

Hoppe, E. I. & Kusterer., D. J. (2011), 'Behavioral biases and cognitive reflection', *Economic Letters* **110**(2), 97–100.

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

Knox, R. E. & Inkster, J. A. (1968), ‘Posdecision dissonance at post time’, *Journal of Personality and Social Psychology* **8**(4), 319–323.

Montgomery, D. W. (1972), ‘Markets in licenses and efficient pollution control programs’, *Journal of Economic Theory* (5), 395–418.

Oechssler, J., Roider, A. & Schmitz, P. W. (2009), ‘Cognitive abilities and behavioral biases’, *Journal of Economic Behavior and Organization* **72**, 147–152.

Phillips, Q. R., Battalio, R. C. & Kogut, C. A. (1991), ‘Sunk and opportunity costs in valuation and bidding’, *Southern Economic Journal* **58**(1), 112–128.

Robalo, P. & Sayagy, R. (2013), ‘Paying is Believing: The Effect of Costly Information on Bayesian Updating’.

Roodhooft, F. & Warlop, L. (1999), ‘On the role of sunk costs and asset specificity in outsourcing decisions: a research note’, *Accounting, Organizations and Society* (24), 363–369.

Staw, B. M. (1976), ‘Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action’, *Organizational Behavior and Human Performance* **16**, 27–44.

Tan, H.-T. & Yates, J. F. (1995), ‘Sunk cost effects: The influences of instructions and future return estimates’, *Organizational Behavior and Human Decision Processes* **63**(3), 311–319.

Thaler, R. (1980), ‘Toward a positive theory of consumer choice’, *Journal of Economic Behavior and Organization* **1**, 39–60.

van Dijk, E. & Zeelenberg, M. (2003), ‘The Discounting of Ambiguous Information in Economic Decision Making’, *Journal of Behavioral Decision Making* **16**, 341–352.

## Appendix A: Tables

Table 1: Experimental parameters

| Treatment | T0 | T100 | T200 |
|---|---|---|---|
| $p_0^A$ (EE) | 0 | 100 | 200 |
| Initial cash (EE) | 900 | 4900 | 8900 |
| $A_0$ (units) | 40 | 0 | 0 |
| $Q$ (units) | 50 | 50 | 50 |
| $p_1^A$ (EE) | uniformly distributed between 50 and 90 | | |
| $p^B$ (EE) | 30 | 30 | 30 |
| $p$ (EE) | 300 | 300 | 300 |

Table 2: Descriptive statistics for the sample who invested

| Treatment | T0 | T100 | T200 | Total |
|---|---|---|---|---|
| N | 48 | 46 | 48 | 142 |
| Economics or Business students | 67% | 57% | 71% | 65% |
| Males | 48% | 54% | 71% | 58% |
| Initial asset called "A" | 52% | 50% | 58% | 54% |
| Consistent answers to the HL questions | 75% | 68% | 75% | 73% |
| Risk aversion (average "safe" lotteries) | 5.38 | 5.43 | 5.29 | 5.37 |
| Average correct answers to the cognitive quiz | 2.67 | 3.13 | 3.12 | 2.97 |

Table 3: Decision variables: all sample

| | N | Use of alternative asset | Sales of the initial asset | Purchases of the initial asset |
|---|---|---|---|---|
| | | (1) | (2) | (3) |
| T0 | 48 | 22.44 (2.74) | 14.21 (2.48) | 1.77 (.52) |
| T100 | 46 | 15.70 (2.43) | 9.11 (2.03) | 3.41 (.63) |
| T200 | 48 | 19.46 (2.54) | 11.65 (2.24) | 2.19 (.57) |

Note: Standard errors in parentheses

Table 4: Averages for the subsample of consistent HL lottery choices

| | N | Use of the alternative asset | Sales of the initial asset | Purchases of the initial asset |
|---|---|---|---|---|
| | | (1) | (2) | (3) |
| T0 | 36 | 26.02 (3.17) | 17.22 (2.93) | 1.19 (0.53) |
| T100 | 31 | 16.90 (2.99) | 10.06 (2.48) | 3.16 (0.8) |
| T200 | 36 | 19.69 (3.03) | 11.92 (2.69) | 2.22 (0.67) |

Note: Standard errors in parentheses

Table 5: Averages of units used from the alternative asset by cognitive level

| | Lowest 75th percentile | | Highest 25th percentile | |
|---|---|---|---|---|
| | N | Units of alternative asset | N | Units of alternative asset |
| T0 | 36 | 16.33 (2.68) | 12 | 40.75 (4.30) |
| T100 | 25 | 14.32 (3.09) | 21 | 17.33 (3.91) |
| T200 | 30 | 16.77 (2.70) | 18 | 23.94 (5.00) |

Note: Standard errors in parentheses

Table 6: Non-parametric test of treatment differences by cognitive level

| Cognitive level | H0: T0=T100 | H0: T0=T200 | H0: T100=T200 |
|---|---|---|---|
| High | 0.0005 | 0.011 | 0.158 |
| Low | 0.269 | 0.362 | 0.171 |

Note: 1-sided $p$-values of the Wilcoxon-Mann-Whitney test are reported

Table 7: The effect of the cognitive ability

| Dependent variable: | Units of the alternative asset | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Constant | 31.38*** | 27.86*** | 30.56*** | 30.47*** |
| | (4.651) | (6.434) | (6.581) | (6.221) |
| Cognitive | 5.183*** | 5.083*** | 8.844*** | |
| | (1.425) | (1.453) | (2.120) | |
| CRT | | | | 10.49*** |
| | | | | (2.160) |
| T100 X Cognitive | | | -7.367** | |
| | | | (2.919) | |
| T200 X Cognitive | | | -4.586 | |
| | | | (3.772) | |
| T100 X CRT | | | | -9.296*** |
| | | | | (3.091) |
| T200 X CRT | | | | -5.889* |
| | | | | (3.487) |
| T100 | -8.239* | -8.330* | -8.734** | -8.511** |
| | (4.423) | (4.487) | (4.368) | (4.286) |
| T200 | -1.949 | -2.230 | -2.930 | -2.670 |
| | (3.923) | (4.113) | (4.127) | (4.032) |
| Time | | 2.267 | 1.573 | 1.282 |
| | | (4.073) | (4.210) | (4.036) |
| Time sq. | | -0.156 | -0.0388 | -0.00632 |
| | | (0.623) | (0.646) | (0.615) |
| Session dummies | Yes | Yes | Yes | Yes |
| Observations | 142 | 142 | 142 | 142 |
| R-squared | 0.148 | 0.157 | 0.185 | 0.209 |

Note: Heteroskedastic robust standard errors in parentheses. The control group T0 is the omitted category. "Cognitive" and "CRT" are the standardized scores for the 5-question cognitive score and the cognitive reflection test score, respectively. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

# Appendix B: Instructions Screens

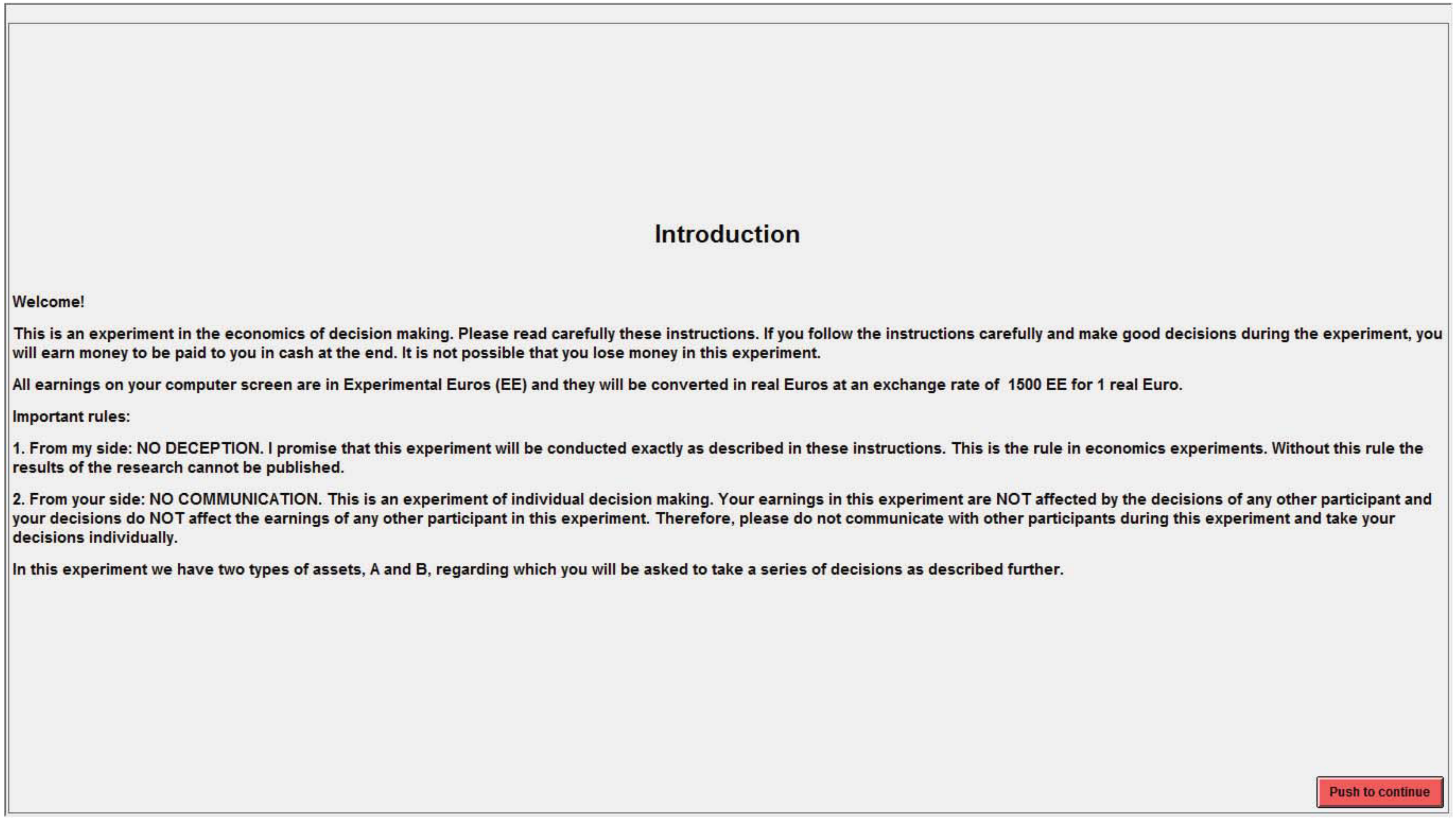


Figure 1: Introduction Screen

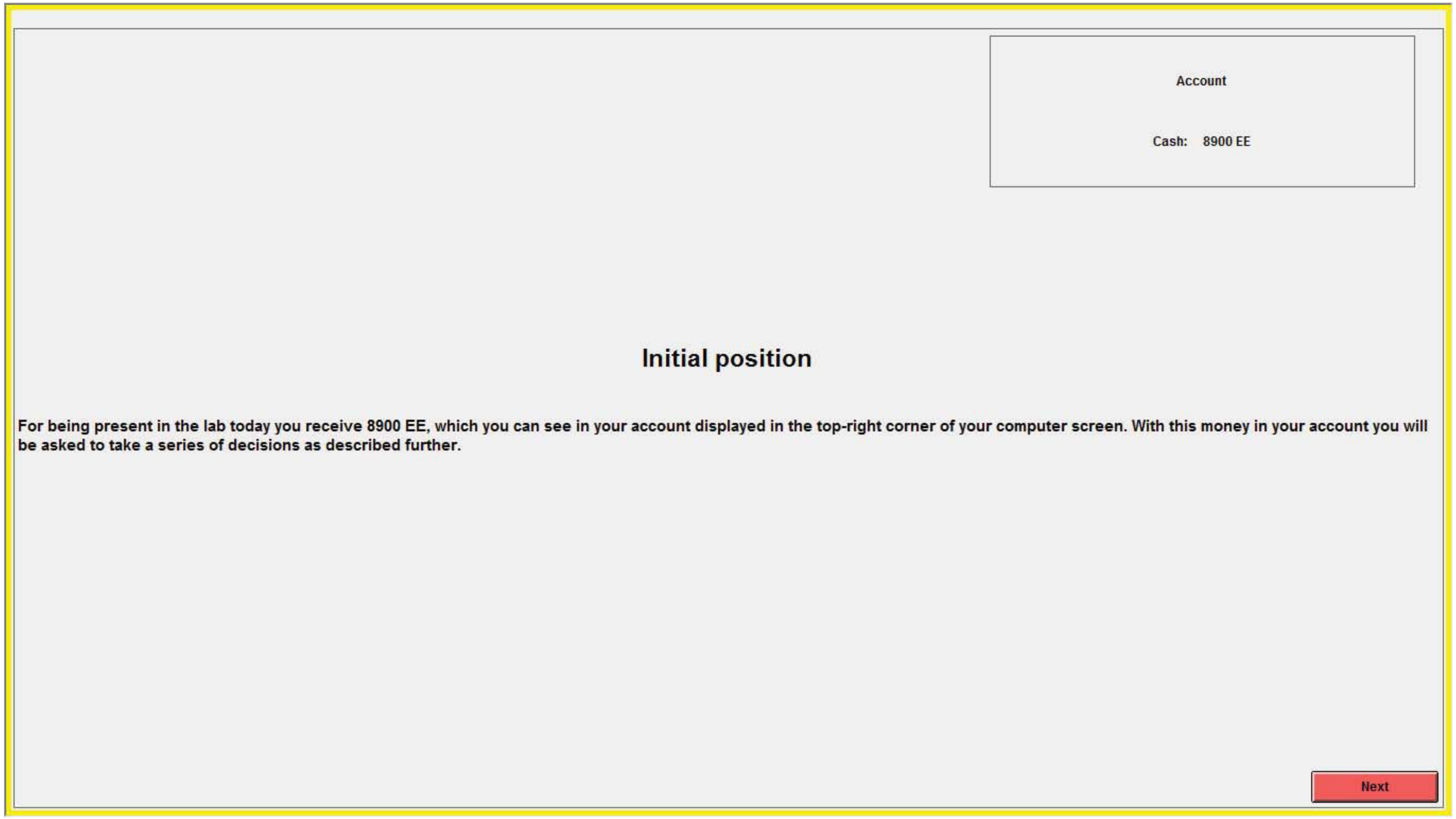


Figure 2: Initial position Screen

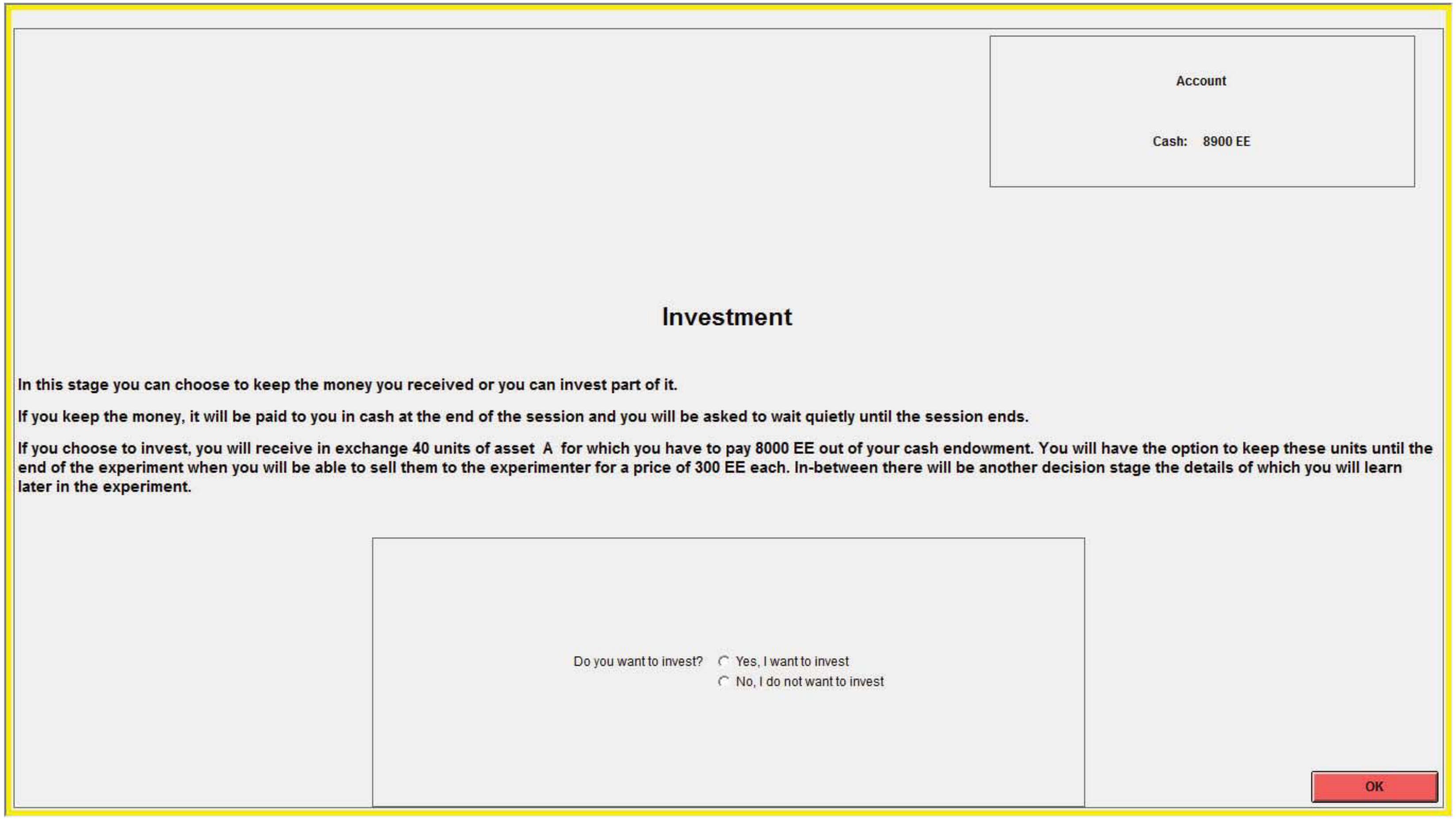


Figure 3: Investment Screen

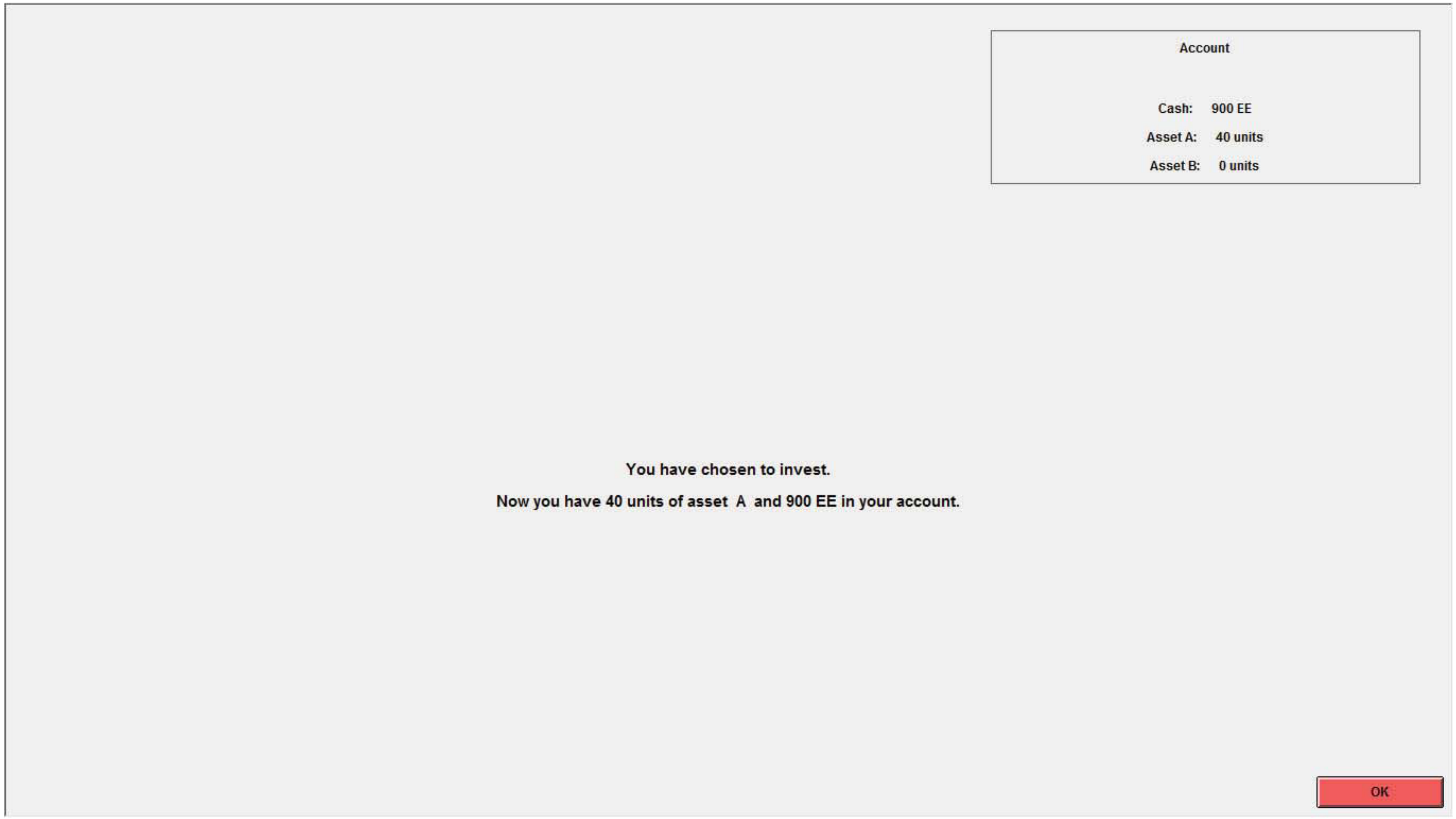


Figure 4: Financial position after investment Screen

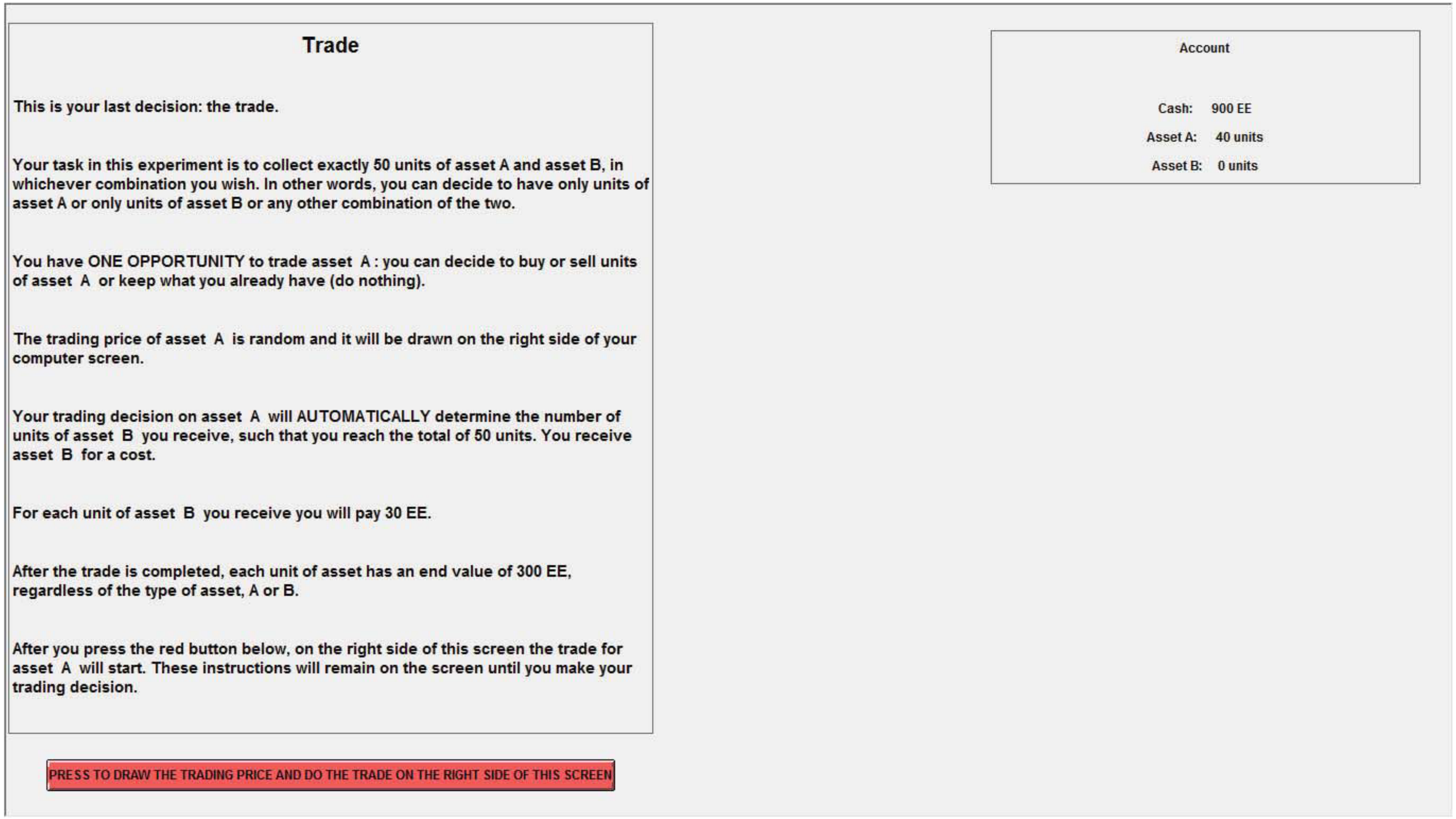


Figure 5: Trade 1 Screen

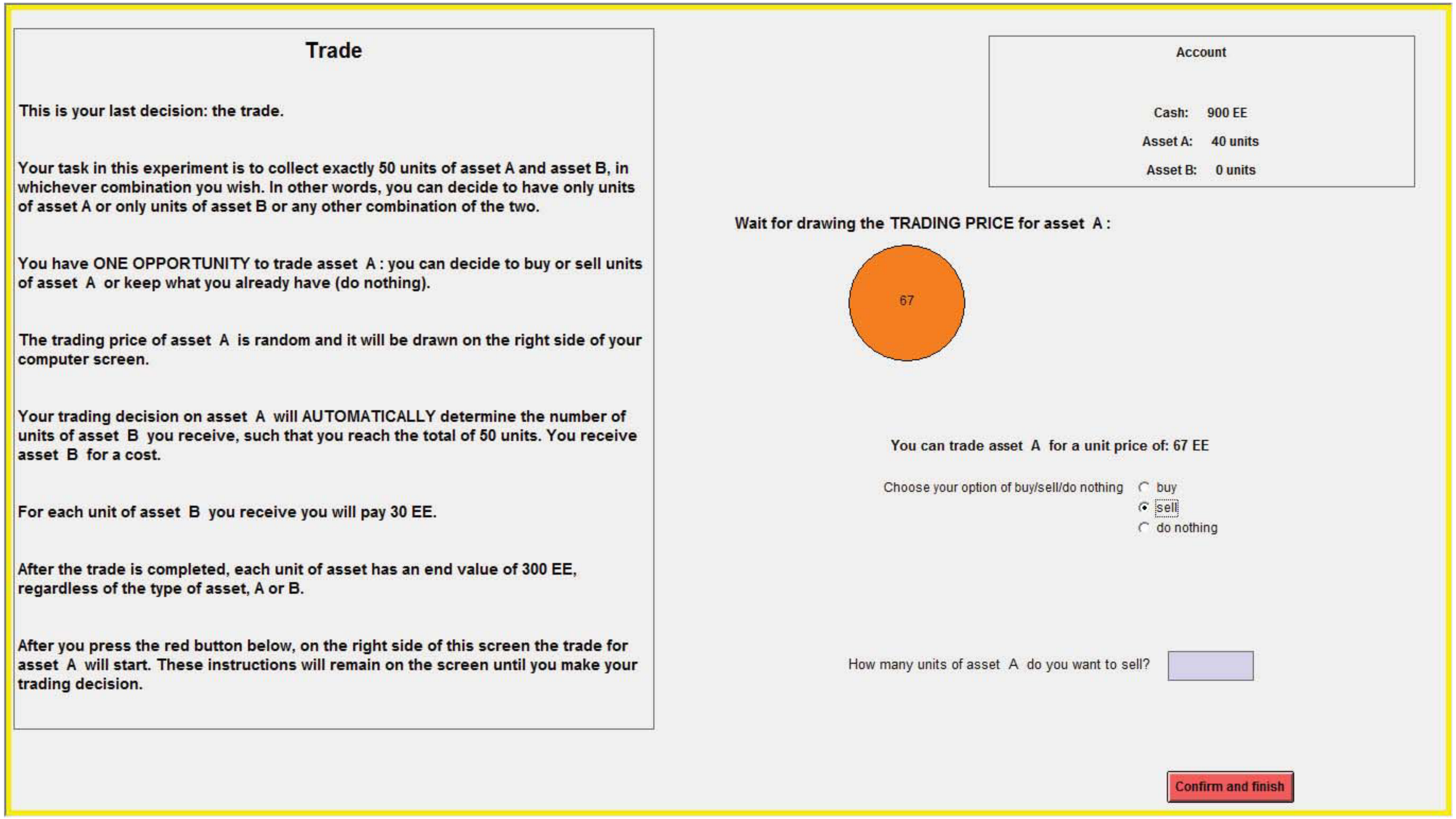


Figure 6: Trade 2 Screen

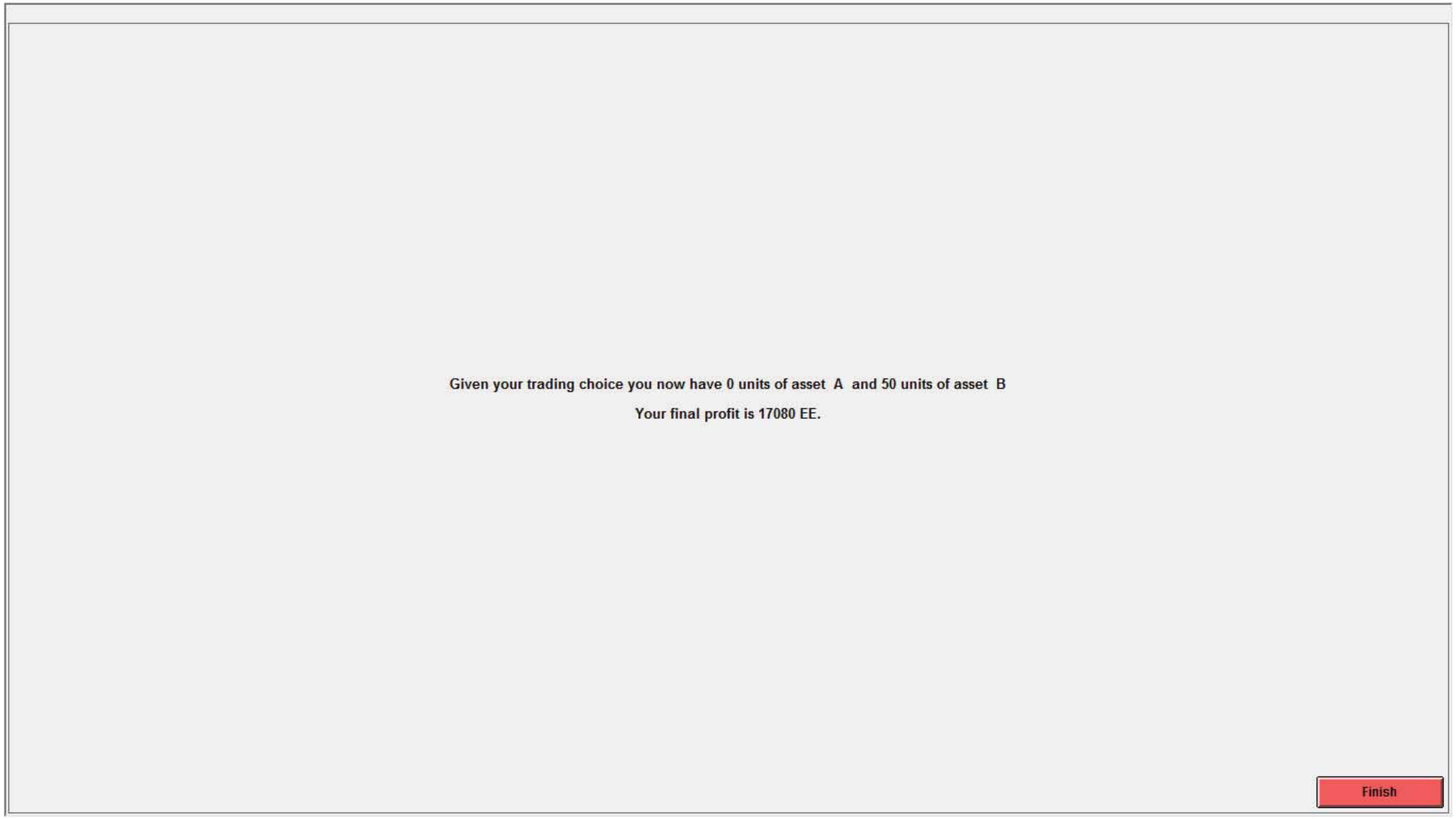


Figure 7: Profit Screen

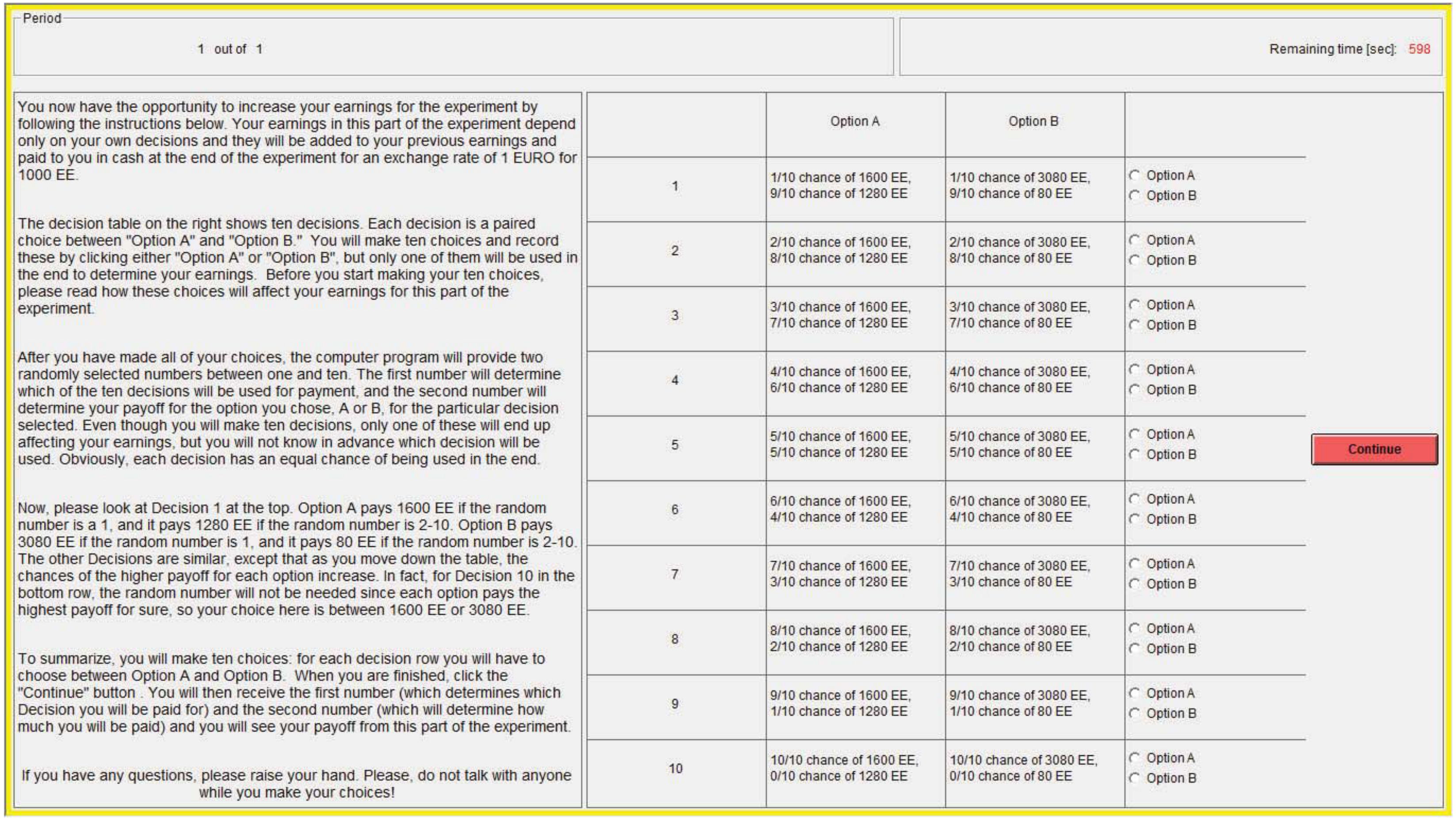


Figure 8: Holt and Laury risk aversion questions

## Appendix C: The Cognitive Quiz

**The CRT questions:**
*Question 1:* An apple and an orange cost \$1.10 in total. The apple costs \$1.00 more than the orange. How much does the orange cost (in \$)?

*Question 2:* If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets (in minutes)?

*Question 3:* In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake (in days)?

**The math questions**
*Question 4:* Half of $-[-a + (b - a)]$ equals:
(A) $a - b/2$
(B) $a + b$
(C) $2a - b$
(D) $2a + 2b$
(E) $-a - b$

*Question 5:* If $x = y - 2$ and $xy = 48$, which of the following CANNOT equal either $x$ or $y$?
(A) 6
(B) 8
(C) 12
(D) -6
(E) -8

## Appendix D: Figures

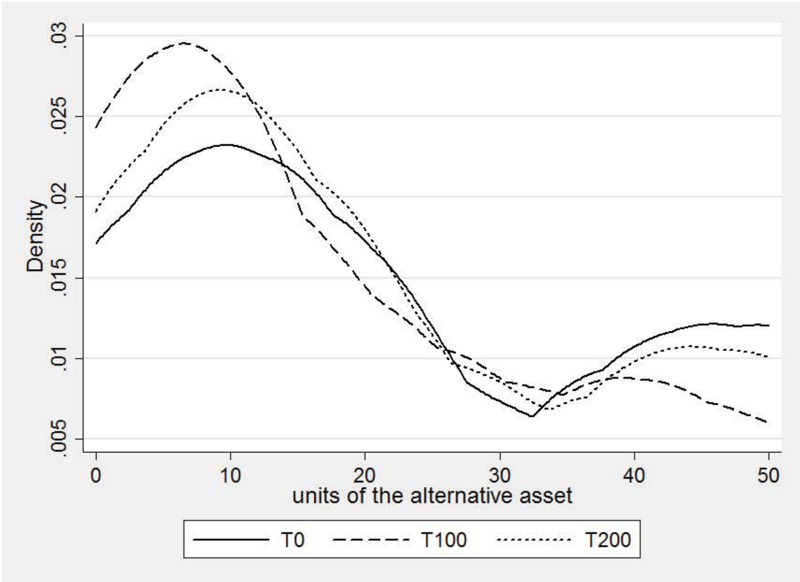


Figure 9: Distribution of units used from the alternative asset: All sample

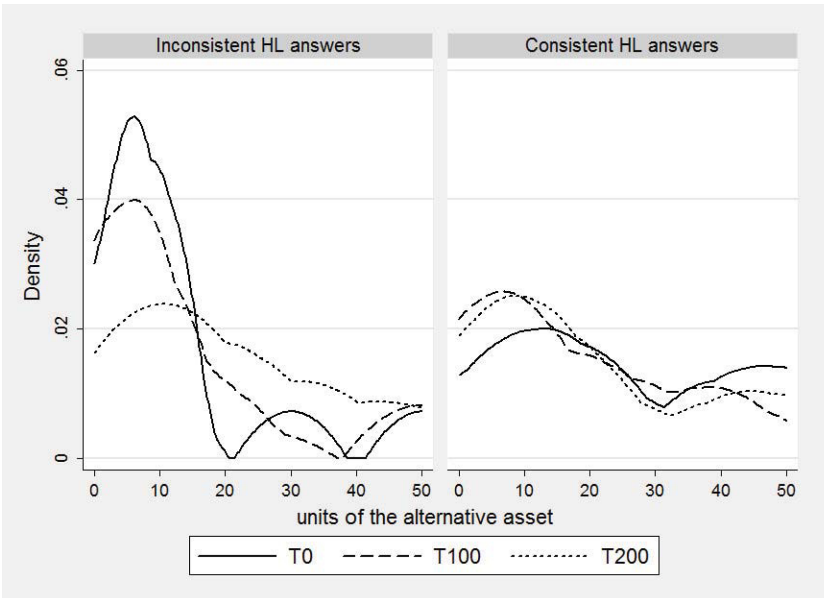


Figure 10: Distribution of the units used from the alternative asset by consistency of the answers to HL lotteries

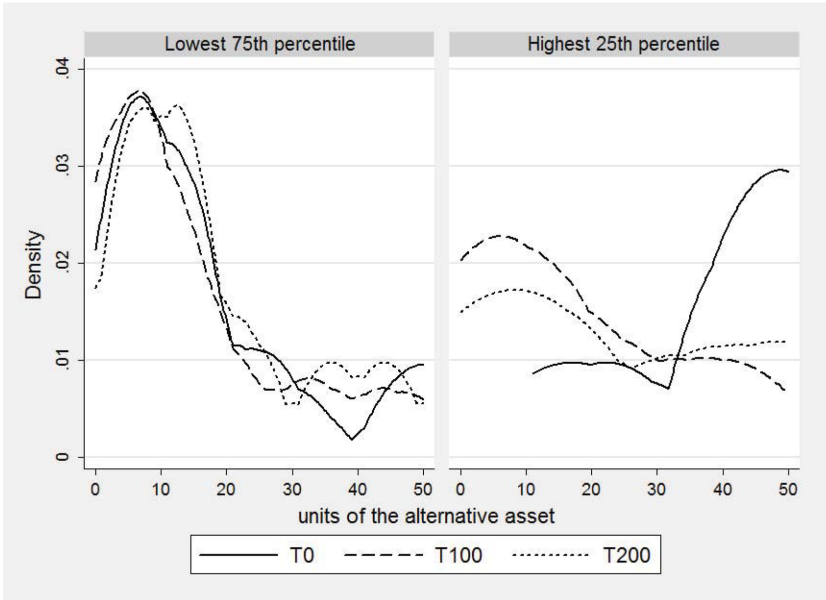


Figure 11: Distribution of the units used from the alternative asset by cognitive score

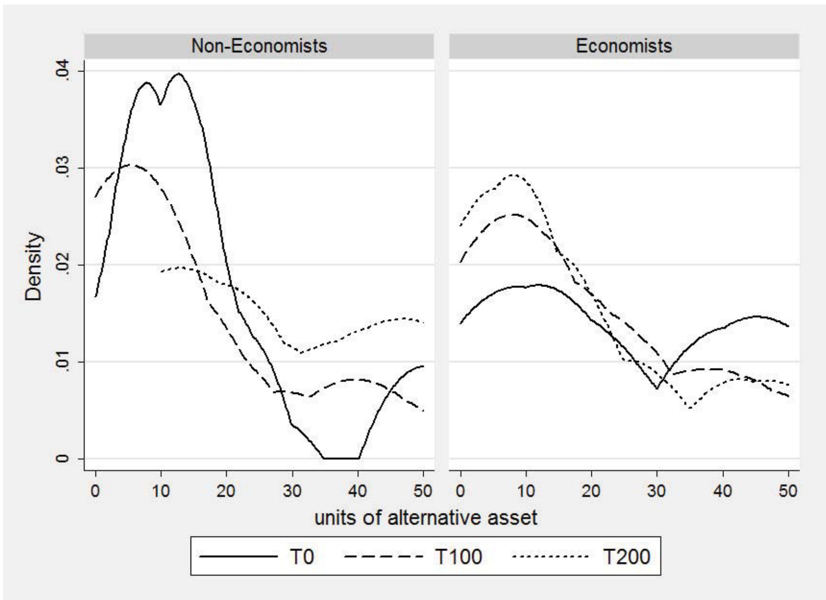


Figure 12: Distribution of the units used from the alternative asset by field of studies