Universidad de Granada
Departamento de Teoría e Historia Económica

Explorations of the Psychological Component in Economic Decisions

Tesis Doctoral

Melanie Parravano Goossens

Director: Dr. Nikolaos Georgantzís

Granada, Abril 2012

Editor: Editorial de la Universidad de Granada
Autor: Melanie Parravano Goossens
D.L.: GR 2881-2012
ISBN: 978-84-9028-220-5

La memoria de tesis titulada “**Explorations of the Psychological Component in Economic Decisions**” que presenta la doctoranda *Dña. Melanie Parravano Goossens* para optar al título de Doctora en Economía, ha sido realizada en el Departamento de Teoría e Historia Económica, dentro del programa de doctorado Empirical Economics (E2), bajo la dirección del *Doctor Nikolaos Georgantzís*.

Granada, Abril de 2012

Doctoranda

Fdo: Dña. Melanie Parravano Goossens

Director

Fdo: D. Nikolaos Georgantzís

Agradecimientos

La materialización de este proyecto no habría sido posible sin el apoyo de muchas personas que compartieron conmigo sus conocimientos, me dieron valiosos consejos y me brindaron su apoyo.

A mi tutor Nikos le agradezco infinitamente, entre tantas otras cosas, porque de principio a fin ha confiado en mi capacidad para culminar con éxito esta empresa, infundiéndome la energía necesaria para lograrlo. Le agradezco de la misma manera, su guía certera, los conocimientos transmitidos, así como su apoyo personal y material (con fondos del Proyecto de Excelencia de la Junta de Andalucía, P07-SEJ-03155).

A Pablo Brañas tengo que agradecer el haber creado y promovido el Ph.D. en *Empirical Economics* y por sus ideas y entusiasmo, que hicieron posible el inicio de uno de los proyectos que conforman esta tesis.

A Julia Morales, siempre jovial y alegre, le agradezco lo mucho que aprendí sobre la práctica de psicología experimental y cognitiva durante los numerosos días que pasamos juntas en los laboratorios de psicología.

Agradezco asimismo a mis demás coautoras, Aurora, Teresa Bajo y Ainhoa por la fructífera colaboración.

Un particular agradecimiento merecen también Luis Enrique, Luis Alejandro y Alexandra por su ayuda en la ejecución de las sesiones experimentales de las que se desprenden los resultados expuestos en los dos primeros ensayos. A Luis Enrique, le agradezco además por su apoyo, por las discusiones que nutrieron esta tesis y las recomendaciones derivadas de su experticia en el análisis de datos.

Estoy agradecida a mis profesores del E^2, Natalia, Tere, Carlos Sánchez, Francis, Tate, Julián, Juliette, Ana, Henry y Betty por ayudarme a ampliar mis conocimientos y ofrecerme las herramientas necesarias para iniciar esta investigación.

También deseo expresar mi agradecimiento a los participantes de los seminarios internos del LEE, del SEET Workshop 2010, del 6to y 7mo IMEBE, del ESHIA Workshop 2010, del IAREP/SABE/ICABEEP Conference 2011 y del ESA European Conference 2011, por sus valiosos comentarios.

Agradezco de manera especial a Matt por su ayuda en la corrección del texto de esta tesis, así como por su inagotable paciencia y apoyo.

A mis compañeros del doctorado Filippos, Antonis, Estrella, Cristian, Bralind, Fady Sascha y en especial a mis amigos Alexandra, Luis Alejandro y Levent por hacer mucho más placentera esta etapa. Y a mis actuales colegas en UEA, Ailko y Axel, por el ánimo que le infundieron a la recta final.

Finalmente, agradezco a mi familia y en especial mis padres a quienes les debo cada uno de mis logros.

Melanie Parravano

Resumen

Esta tesis explora empíricamente distintos aspectos de las decisiones individuales en contextos estratégicos y de riesgo. Para ello se utiliza un enfoque interdisciplinario que combina la Psicología y la Economía Experimental. La tesis se compone de tres ensayos independientes. Los dos primeros, tienen como objetivo común explicar, sobre la base de las capacidades cognitivas individuales, resultados que contradicen sistemáticamente las predicciones de la teoría de juegos clásica. En el tercer ensayo, por su parte, se revisan las características de una tarea específicamente diseñada para *elicitar* las preferencias hacia el riesgo.

En particular, en el primer ensayo nos dedicamos a estudiar en el laboratorio las decisiones en el *Traveler´s Dilemma*. Este juego plantea una paradoja, ya que cambios en un parámetro que no altera el equilibrio de Nash causan cambios sustanciales en las decisiones. En particular, cuando el valor del parámetro de castigo es bajo, los individuos escogen mayoritariamente estrategias que se aleja de las predicción de equilibrio de Nash, mientras que cuando el parámetro de castigo es alto, las decisiones convergen al equilibrio de Nash. Además de reproducir los resultados encontrados en implementaciones experimentales previas, encontramos que las habilidades cognitivas individuales, en particular la capacidad de razonamiento y la memoria de trabajo, ayudan a explicar las elecciones individuales en este contexto estratégico. Sin embargo, contrario a anteriores conjeturas que atribuyen este cambio de comportamiento a errores derivados de fallos cognitivos, encontramos que los individuos con mayor capacidad de razonamiento y memoria de trabajo son más sensibles al cambio en este parámetro. También encontramos que los individuos con mayor capacidad cognitiva, son más sofisticados desde el punto de vista estratégico, ya que utilizan una estrategia de mínimo recorte que proporciona mayores beneficios que la estrategia de Nash.

Utilizando un enfoque similar, en el segundo ensayo presentamos los resultados de un *Matching Pennies* Asimétrico, en el cual la teoría sugiere que ambos jugadores han de utilizar estrategias mixtas, es decir, aleatorizar sus decisiones con la finalidad de ser impredecibles. Sin embargo, se ha documentado que ante la presencia de una asimetría favorable, los individuos tienden a escoger con una alta frecuencia las estrategias asociadas a elevados pagos potenciales, aun cuando esto los hace predecibles y por tanto

estratégicamente vulnerables a sus oponentes. Nuestros resultados confirman fuertes desviaciones del equilibrio en estrategias mixtas. Más aun, mostramos que esta atracción hacia las estrategias que conllevan la posibilidad de altos pagos, se asocian a menores niveles de inhibición y capacidad de razonamiento. Encontramos adicionalmente que individuos con una mayor memoria de trabajo tienen una mayor probabilidad de explotar este comportamiento en sus oponentes, lo cual les permite incrementar su ganancia esperada.

En el tercer ensayo proponemos una tarea simple para *elicitar* las actitudes frente al riesgo. La *SGG lottery-panel task* que consiste en una serie de loterías construidas compensando opciones más riesgosas con un mayor retorno al riesgo. Utilizando la técnica de Análisis de Componentes Principales identificamos endógenamente dos dimensiones del comportamiento individual en contextos de riesgo: la disposición promedio a tomar riesgo y la sensibilidad a variaciones en el retorno al riesgo. Reportamos resultados derivados de la implementación de la tarea a una amplia muestra de sujetos y discutimos las regularidades y los beneficios de capturar estas dos dimensiones tanto para caracterizar el comportamiento bajo riesgo como para explicar las decisiones en otros contextos. Finalmente, extendemos el espectro de la tarea original al proponer tres nuevos tratamientos que combinan premios elevados y la posibilidad de pérdidas (loterías mixtas). Mostramos como estos nuevos tratamientos capturan regularidades en el comportamiento bajo riesgo y conservan como propiedad la capacidad de capturar ambas dimensiones de las actitudes frente al riesgo.

Contents

1. Introduction

This thesis explores empirically different aspects of individual decision making in strategic and risky contexts. For this we use an interdisciplinary approach that combines Psychology and Experimental Economics. The thesis is composed by three independent essays. The first two essays combine Experimental Game Theory with Cognitive Psychology. In these studies the common objective is to explain behavior and paradoxical findings reported elsewhere on the basis of individual cognitive skills. The third essay is related to individual decision making under uncertainty and reviews the features of a specific task designed to elicit risk preferences.

In particular, in the first essay entitled “Cognitive Underpinnings of Behavior in the Traveler’s Dilemma: Treasures and Contradictions Revisited”, we focus on the well-known paradoxical shift from equilibrium to non-equilibrium play in the Traveler’s Dilemma Game, following a change in the theoretically irrelevant “punishment” parameter. Apart from reproducing previous results in the literature, we show that cognitive skills, in particular reasoning ability and working memory, play a role in the choices observed in this game. Against previous conjectures that the shift in behavior that occurs when the theoretically irrelevant “punishment” parameter goes from high to low might be due to poor understanding of the game, we find that this apparently paradoxical behavior is observed more often among subjects with a higher cognitive skills. We also find that more cognitively sophisticated individuals are indeed more strategically sophisticated players, which does not mean that they play the Nash equilibrium when the punishment/reward parameter is low, but that they might anticipate modal behavior and best respond to other players’ actions, using a minimal undercutting strategy.

Using a similar approach, the second essay entitled “On the Cognitive Foundations of Mixed Strategy Equilibrium Failure in the Lab” deals with an Asymmetric Matching Pennies game, in which theory expects subjects’ strategies to be randomized in a way which makes them impossible to predict. The paradoxical finding is that subjects are systematically attracted by high own-payoffs choices even when such behavior makes them predictable and, thus, strategically vulnerable to their opponents. Our results confirm strong deviations from the unique mixed strategy Nash equilibrium predictions.

Moreover, we show that this unprofitable attraction is associated with low degrees of inhibition and reasoning ability, as measured in specific psychometric tests. We additionally find that player with higher working memory are more likely to anticipate this behavior, thus achieving higher payoffs. We also find gender differences and playing order effects, due possibly to introspection.

In the third essay, entitled “The Lottery-Panel Task for Bi-dimensional, Parameter-Free Elicitation of Risk Attitudes: Implementation and Results”, we propose a simple task for the eliciting attitudes toward risky choice. The SGG lottery-panel task, which consists of a series of lotteries constructed to compensate riskier options with higher risk-return trade-offs. Using the Principal Component Analysis technique we endogenously identify two main components of individual behavior in risky contexts: a subject’s average willingness to choose risky prospects and their sensitivity towards variations in the return to risk. We report results from a large dataset obtained from the implementation of the SGG lottery-panel task and discuss regularities and the desirability of its bi-dimensionality both for describing behavior under uncertainty and explaining behavior in other contexts. Finally, we extend the scope of the original task by proposing three new treatments where the high stakes and mixed (gains-losses) outcomes are introduced. Results from implementing these new treatments capture several regularities in risk taking related to stakes size and domain effect. The new treatments also capture two desirable dimensions of risk attitudes: a subjects´ overall risk taking and their response to changes in the return to risk.

The boundaries between Economics and Psychology are not only an underexplored field, but, also a very fruitful one for social scientists altogether. This thesis is an example of how the Psychologists' work on cognitive inventories can assist the Economist in explaining behavior in strategic interaction and a proposal on how the economist should develop instruments not just for theory testing, but also for capturing idiosyncratic decision making features which could explain behavior in a variety of contexts.

Essay 1

2. Cognitive Underpinnings of Behavior in the Traveler's Dilemma: Treasures and Contradictions Revisited[1]

2.1 Introduction

There is abundant experimental evidence showing that individual behavior frequently deviates from standard game theory equilibrium predictions. This is the case of the Traveler's Dilemma (Basu, 1994), a game whose equilibrium prediction often fails to organize observed behavior in the lab, which in turn, can be easily explained with intuitive reasoning.

A synthesized version of the parable associated with the game is the following. Two travelers returning home from an island, where they bought identical antiques, discover that the airline has smashed them. To compensate for damages, the airline manager, who does not know the real value of the antique, offers the following scheme. Each of the two travelers has to independently make a claim, between c^1 and c^2, for the value of the antique. If both claim the same amount, that is, $c_i=c_j$, then it is reasonable to assume that they are telling the truth and so each of them will be reimbursed c_i (or c_j) units of money. However if traveler *i (j)* claims a larger amount than traveler *j (i)*, then the manager will treat the lower claim, that is, c_j (c_i), as the true cost. Additionally, to discourage false claims, he will charge a penalty "*R*" to the higher claimant that will be transferred as a reward to the lower claimant. That is, the lower claimant *j (i)* will receive c_j+R (c_i+R) while the higher claimant *i (j)* will receive c_j-R (c_i-R).

All standard game theoretic solution concepts predict that the unique equilibrium is that both players select the lowest claim, which is $c_i=c_j=\underline{c}$. Thus, the unique equilibrium, the only Nash equilibrium and only rationalizable outcome is ($\underline{c}$,$\underline{c}$) despite the real value of

[1] An earlier version of this manuscript has been presented at: SEET Workshop in Marrakech, the 6th IMEBE in Bilbao, the ESHIA Workshop in Alessandria, the IAREP / SABE / ICABEEP Conference in Exeter and the ESA European Conference in Luxembourg. The coauthors of that version and future ongoing extensions: M. Teresa Bajo-Molina, Pablo Brañas-Garza, Nikolaos Georgantzís, and Julia Morales are gratefully acknowledged. Also financial support by the Junta de Andalucía (P07-SEJ-03155) and the Spanish Ministry of Science and Innovation (ECO2008-04636/ECON; ECO 2010-17049) is gratefully acknowledged.

the object or the size of *R*. Nevertheless, if the penalty for choosing the higher claim is not severe, the consequence of a unilateral deviation from the equilibrium strategy is a small monetary loss compared to the (substantially) higher payoffs that can be achieved by (extreme) bilateral deviations from equilibrium. Therefore, as Basu (1994) noted “it seems very unlikely that any two individuals, no matter how rational they are and how certain they are about each other's rationality, each other's knowledge of each other's rationality, and so on, will play *($\underline{c}$,$\underline{c}$)*. It is likely that each will play a large number in the belief that so will the other and thereby they will both get large payoffs” (p. 392). Indeed, Basu’s intuition of large claims being chosen when the penalty is low has been widely confirmed by experimental implementations for both repeated (Capra, Goeree, Gomez, & Holt, 1999) and one-shot (Cabrera, Capra, & Gómez, 2007; Goeree & Holt, 2001; Rubinstein, 2007) Traveler’s Dilemma games.

In addition, the failure of the Nash equilibrium to organize the data obtained in various Traveler’s Dilemma experiments is further amplified by the fact that changes in the theoretically irrelevant penalty/reward parameter (*R*) strongly affect players’ strategies. To be precise, there is a strong negative correlation between level of the *R* parameter and average claim levels, i.e. when the *R* parameter is sufficiently high, behavior tends to conform closely to the Nash equilibrium, but claims rise to the maximum level as the *R* parameter approaches zero (Capra et al., 1999). This finding was corroborated by (Goeree & Holt, 2001) in what they coined as a “*treasure*” and a “*contradiction*” of game theory[2].

Several conjectures have been proposed to address both the Traveler’s Dilemma (henceforth TD) paradox and the effect of the *R* parameter on claim levels. In particular, regarding the deviations from the Nash equilibrium in low penalty TD we can distinguish two alternative explanations, namely: 1) Cognitive bottlenecks impede individuals to perform all the reasoning steps necessary to deduce that the only rational strategy is to claim the minimum; 2) People understand that strongly deviating from the Nash equilibrium it is possible to obtain much higher payoffs and expect the other player to realize the same.

[2] Basu, Becchetti, and Stanca (2011) using also extreme (high-low) values of R, explored the effect of asymmetric changes in this parameter, and found that subjects were highly more sensitive to a change in their own reward-penalty, than that of their co-player.

The first conjecture implies a relaxation of orthodox game theory in which players are supposed to be able to perform indefinitely recursive reasoning (Colman, 2003). In fact cognitive limitations are one of the main assumptions of most behavioral game theory models, such as the Cognitive Hierarchy Model (Camerer, Ho, & Chong, 2004), the level-k models (Costa-Gomes & Crawford, 2006; Nagel, 1995; Stahl & Wilson, 1995) and the Quantal Response Equilibrium model (McKelvey & Palfrey, 1995) and its logit formulation (Anderson, Goeree, & Holt, 2002), which have been strongly influenced by data from experiments, for instance the TD game[3]. The second explanation, on the other hand, is closer to conventional decision theory, in which, it is assumed that rational agents always choose those alternatives that maximize their payoffs relative to their beliefs.

According to Goeree and Holt (2001), results in the low *R* (contradiction) treatment could be due to the fact that deletion process necessary to eliminate all the alternatives that are never a best response may be too lengthy for human subjects with limited cognitive abilities. Nevertheless, the argument is in conflict to the evidence of another experimental implementation of the TD by Becker, Carter, and Naeve (2005), where game theory experts were asked to submit both a strategy and their belief concerning the average strategy of their opponents. Surprisingly, they found that less than 6% played the Nash equilibrium, while almost 20% played the maximum claim[4]. Therefore, this implementation gives evidence against both the misunderstanding of the strategic nature of the game and the incapacity to deduce the rational equilibrium as main explanations to deviations from the Nash prediction in the TD.

Recent research on the TD game has moved away from characterizing and explaining modal behavior, focusing instead on the issue of behavior heterogeneity (Basu, Becchetti, & Stanca, 2011; Brañas-Garza, Espinosa, & Rey-Biel, 2011; Rubinstein, 2006, 2007), which can be caused among other reasons by cognitive ability differences. On one hand, Rubinstein (2007) implemented among a large sample of subjects a one-

---

[3] In addition, we find in the literature several theoretical approaches to model the TD paradox. Among those for one-shot interactions of this game we can mention: Noisy introspection models (Cabrera et al., 2007; Goeree & Holt, 2004), heterogeneous social preferences (Erlei, 2008), ambiguity about other players behavior (Eichberger & Kelsey, 2011) and iterated regret minimization (Halpern & Pass, 2012).

[4] Moreover, from their data they estimated that the best response to the average strategy was to play the pure strategy 97 (only 3 units under the highest possible claim of 100).

shot, low R, hypothetical payoff TD[5] obtaining claims and associated response times, in order to check the hypothesis that fast decisions should reflect instinctive choices whereas longer decision times should be associated with strategic reasoning. His results are in line with the preceding hypothesis, in the sense that choices corresponding to the maximum, instinctive, claim (300) were associated with significantly shorter average reaction times than minimal undercutting *strategic choices* in the range 295-299. On the other hand, Brañas-Garza, Espinosa and Rey-Biel (2011), used participants' self reported justification of choices (strategies) and demonstrated that individuals' alleged motivations are useful to predict their choices. Moreover they found coherence between those motivations and independent measures of behavior in other experimental tasks and ability measures. Finally, Basu et al. (2011) define three strategic types based on participants' ex-post descriptions of their strategies (Nash or individually rational, team strategic and irrational) and find heterogeneity of players' preferences is supported by observed claim-belief pairs and self-revealed strategies at the end of the game.

This essay contributes to the existing literature on the Traveler's Dilemma game by exploring the cognitive foundations of paradoxical behavior and behavioral heterogeneity in this particular strategic context. This essay also contributes to the recent literature exploring the link between cognitive abilities and behavior in strategic interaction contexts (Brañas-Garza, García-Muñoz, & Hernan, 2011; Burnham, Cesarini, Johannesson, Lichtenstein, & Wallace, 2009; Devetag & Warglien, 2003; Rydval, Ortmann, & Ostatnicky, 2009).

Our approach consists of exogenously measuring player's potentially relevant cognitive abilities and their decisions in the game. For this purpose, we asked 81 participants to play one shot TDs, using two extreme *R* values, like in Goeree and Holt (2001). Previous to the experiment, we assessed different aspects of cognition of the participants, such as analytic intelligence, working memory capacity, flexibility of thinking and response inhibition.

Results from our study reveal that cognitive ability differences among individuals, specifically reasoning ability and working memory, help to explain not only general paradoxical findings across the *contradiction* and the *treasure* treatment, but also

[5] Claims in the range 180-300 and R=5, similar to Goeree and Holt (2001) as well as our own implementation of the game.

behavioral heterogeneity within treatments. Our most remarkable result is that cognitive abilities matter, but not in the traditional view that players with higher cognitive capacities will play more in accordance to the theoretical prediction, but that on the contrary, the higher a player's cognitive ability is, the more he will play according to the intuitively expected pattern.

The rest of the paper is organized as follows. Section 2.2 provides a short explanation of the assessed cognitive abilities, and why they might predict behavior in the TD. Section 2.3 illustrates the experimental procedure for both the TD game implementation and cognitive abilities assessment. Section 2.4 is dedicated to describe our dataset. In section 2.5 we show results. Section 2.6 offers conclusions.

2.2 Cognitive abilities and strategic behavior

We considered four potentially relevant aspects of cognition as candidates for explaining behavior in the Traveler's Dilemma game: Non-verbal reasoning ability, working memory, flexibility of thinking, and response inhibition.

2.2.1 Non-verbal reasoning ability

Since in the TD game, participants are confronted with a problem that is completely new to them, they have to rely on their reasoning ability (analytic intelligence) to understand the strategic nature of the game, moreover, we expect that low levels of analytic intelligence[6] can lead to reasoning errors, confusion and the possible use of simple rules such as pick the average. Therefore, we considered that a measure of reasoning ability might be a good predictor of behavior in the TD.

For assessing analytic intelligence, we selected a classical non verbal reasoning test, commonly used for this purpose, the Raven Standard Progressive Matrices test by Raven (1960). Rigas, Carling, and Brehmer (2002) note that performance in Raven's

[6] Carpenter, Just, and Shell (1990) use the term analytic intelligence "to refer to the ability to reason and solve problems involving new information, without relying extensively on an explicit base of declarative knowledge derived from either schooling or previous experience […] Thus, analytic intelligence refers to the ability to deal with novelty, to adapt one's thinking to a new cognitive problem" (p. 405). According to Kyllonen and Christal, (1990) "research has shown that this factor (reasoning ability, i.e. analytic intelligence) is one of the primary determinants if not the primary determinant of the degree to which a person benefits from instruction" (p. 392).

(1976) Advanced Progressive Matrices correlate with performance in two dynamic decision tasks. More close to our study Burnham et al. (2009) found a negative correlation between scores in a short standard psychometric test[7] and entries in a beauty contest game (Nagel, 1995) as well as more dominance violations amongst subjects within the lowest two deciles of cognitive abilities. In contrast, Brañas-Garza et al. (2011), do not find any correlation between beauty contest game entries and scores in the Raven test, but they find that subjects with higher scores in the Cognitive Reflection Test (Frederick, 2005) are more prone to play according to the Nash equilibrium prediction.

2.2.2 Working memory

Apart from requiring general reasoning skills, a shared characteristic of dominance solvable games, such as the TD, is that for deducing the "rational" outcome solution one needs to be capable of performing some level of complex iterative reasoning (I think, you think (you think, I think (…)). According to Goeree and Holt (2004) for one shot games, where learning is not possible but only introspection "this type of iterated reasoning corresponds to considering a sequence of best responses to best responses" (p. 366). Indeed, as we mentioned before, individuals' limitations in the capacity to perform such a process, is one of the main arguments to explain deviations from equilibrium play (Goeree & Holt, 2001; Palacios-Huerta & Volij, 2009).

Coincidently, in cognitive psychology, the term Working Memory (WM), is defined by Baddeley (1992) as "a brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language comprehension, learning and reasoning" (p. 556)[8]. In practical terms WM capacity is reflected in the ability to perform tasks that require several steps with in-between results that need to be kept in mind provisionally. However, humans have limitations in their WM capacity, a bottleneck that restricts their ability to perform such type of tasks and therefore this limitation should also be reflected in the strategic reasoning capacity. For

[7] The test used was developed by the Swedish psychometric company Psykologiförlaget (Sjöberg, Sjöberg, & Forssén, 2006).

[8] According to Conway, Kane, and Engle (2003) "In the past decade, cognitive scientists have entertained the notion that working memory capacity is the 'Factor X' that underlies individual differences in general intelligence" (p. 547). However, more recent evidence reveals that working memory capacity and general intelligence are indeed highly related, but not identical (Conway et. al, 2003).

these reasons, we consider that WM capacity measures should correlate with decisions in the TD. They should be helpful to explain behavioral heterogeneity as well as helping to unravel the TD paradox, since they can provide evidence in favor or against the conjecture that individuals' failure to play the Nash equilibrium (in the low R TD) is due to cognitive bottlenecks that prevent them to perform the iterated deletion of dominated strategies.

Although behavioral game theory models address the issue of limited WM capacity as a possible explanation of individuals' failure to reach the Nash equilibrium and there is broad empirical evidence supporting that WM capacity measures reliably predicts performance in a wide variety of real world cognitive tasks (Engle, Tuholski, Laughlin, & Conway, 1999), the empirical literature linking direct measures of WM and performance in strategic decisions contexts is rather scarce. Some exceptions are a study by Rydval et al. (2009) who found that reasoning errors in three minimalist dominance-solvable guessing games were associated with lower performance on a test of working memory[9]. And, Devetag and Warlien (2003) who found the presence of a significant and positive correlation between subjects' short-term memory score and conformity to standard game-theoretic prescriptions in three games that required iterated reasoning to solve the equilibrium best response[10]**.**

We used two tasks in order to measure different aspects of WM. First a counting span task[11] (Pickering, Baqués, & Gathercole, 1999) which simultaneously demands to store and manipulate information. Secondly, a more complex WM measure, the *N*-Back Task (Kirchner, 1958), which, according to Owen, McMillan, Laird, and Bullmore (2005) "requires on-line monitoring, updating, and manipulation of remembered information and is therefore assumed to place great demands on a number of key processes within working memory" (p. 47).

[9] Specifically they use an "operation span" test (Turner & Engle, 1989). They also found that intrinsic motivation and premeditation attitude were relevant.

[10] Notice that short term memory is related but not equal to WM. WM involves not just keeping in mind information but also doing it while one performs another cognitive process.

[11] There is a broad variety of span tasks aimed to tap different aspects of WM (verbal, operational and counting, among others).

2.2.3 Flexibility of thinking

Executive functions are assumed to serve as cognitive control processes, in particular for planning and organizing behavior. Impairments in these executive functions are often assessed with the Wisconsin Card Sorting Test (WCST). The WCST is generally used to measure the capacity to deduce concepts and to apply a strategy to adapt behavior to changing conditions (Eling, Derckx, & Maes, 2008). Berg (1948) showed that, using the WCST procedure, both learning a rule and shifting to a new rule can be measured simply and adequately, a capacity she referred as flexibility in thinking. Although we didn't have specific priors on how the degree of "cognitive flexibility" could affect subject's decisions in the TD, we considered relevant to explore this dimension of cognition and its possible implications.

2.2.3 Response inhibition

We considered as argued by Dempster (1991) that "although much of the evidence is preliminary, particularly as it applies to text processing and reasoning, it is sufficiently provocative to suggest that inhibitory processes are a neglected, but critically important dimension of intelligence. Inhibitory processes appear to define a basic cognitive dimension that enters into a broad spectrum of intellectual phenomena." (p. 167).

Specifically, response inhibition refers, according to Verbruggen and Logan (2008), to "the suppression of actions that are no longer required or are inappropriate, which supports flexible and goal-directed behavior in ever changing environments" (p. 418). Therefore, we decided to apply the Stop-Signal paradigm, which is very well suited for the study of response inhibition in a laboratory setting.

Table 2.1

Summary of cognitive abilities and related tasks

Cognitive ability	Task
Non-verbal reasoning	Raven Progressive Matrices
Working Memory	Dot Counting Span *N*-Back
Flexibility of thinking	Wisconsin Card Sorting
Response inhibition	Stop-Signal

2.3 Method

2.3.1 Experimental design

The experimental design consists of two different stages. In the first stage each participant individually completed the 5 cognitive tasks (Table 2.1). A detailed description of the tasks implementation is available in the Appendix (5.1.1).

In the second stage, we implemented a standard TD game. In order to facilitate comparability, we used the same set of parameters reported in Goeree's and Holt's (2001) treasures and contradictions paper. In the game, two participants have to choose an integer between 180 and 300, without communicating with each other. If both choose the same number (make identical claims) both receive the amount of money equivalent to the number they claimed[12]. If the numbers chosen are different, both participants earn an amount of money equivalent to the minimum of both quantities. In addition, the participant who has chosen the higher number has to pay a "penalty" of *R* units. These *R* units are received as a reward by the participant who has chosen the lower number (instructions available upon request). Following Goeree and Holt (2001) we used two extreme values for *R*, namely 180 and 5, which correspond to the *treasure* and *contradiction* treatment, respectively.

[12] In our implementation each experimental unit was equivalent to 5 cents of Euro (for example, a claim of 300 was equivalent to 15 Euros).

We first implemented a between-subjects design, that is, half of the participants (40) played the low *R* treatment (R_L) and the other half (41) played the high *R* treatment (R_H). Additionally, to increase the number of observations and allow analysis across subject's decisions, we implemented a within-subjects design. That is, after submitting strategies for the initial part, all participants played the two treatments in random order. Hence, we endogenously obtained four treatment combinations (R_H–R_H –R_L , R_H– R_L–R_H , R_L–R_H –R_L and R_L– R_L–R_H) which were played similarly often (22, 19, 21 and 19 participants, respectively) as a result of randomization. Therefore, for all participants we have their scores for the corresponding five cognitive tasks and three decisions in the TD, as well as information about their age and gender.

2.3.2 Participants

A total of 81 subjects participated in our study. All undergraduate first and second year economics and business students from the University of Granada[13]; (53 females and 28 males) aged on average 19.18 years (S.D.=1.55, age range=18-27). The recruiting process comprehended the following phases: 1) Students from first year courses received an invitation to participate in the project; 2) Interested students provided contact details and time availability for further contact. 3) Potential participants were contacted by email, and a session schedule was proposed, 4) After reaching an agreement, they were informed on the definitive session schedule, either by email or phone.

2.3.3 Procedures

The cognitive tasks were administered at the University of Granada Experimental Psychology Lab, and for completing them participants had to assist to two sessions[14], which lasted around 45 minutes each. The administration of the five tasks to the total of participants was done during a period of five weeks. The order in which tasks were completed by the participants on each session was counterbalanced so that each task was run in each position (first, second, third, fourth and fifth) equally often. Participants

[13] To the best of our knowledge none of them had any prior formal training in game theory.
[14] The two sessions took place in consecutive days.

completed the computerized tasks (Raven Standard Progressive Matrices, *N*-Back and Stop-Signal) individually in rooms with one or two participants, monitored by one experimenter, while tasks that required to be directly administered by the experimenter (Dot Counting Span Task and Wisconsin Card Sorting Task) were performed in a separate room. Each participant received a flat fee of 10 Euros for his/her participation in this part.

After all participants completed the cognitive ability assessment, three consecutive experimental sessions were conducted in the Economics and Business Faculty. Sessions took place the same day[15], in order to minimize the possibility that participants in different experimental sessions might share relevant information that could affect their decisions. The experiment was administered in a paper-and-pencil format and lasted around 50 minutes.

Each experimental session started with the participants receiving a sheet with the instructions for a TD game (choices between 180 and 300). Half of the participants had a small reward/penalty parameter (5) and the other half a high reward/penalty parameter (180). Apart from the instructions, the game was explained to them in a neutral language and avoiding mentioning the amount of the penalty. They were also informed that they would be randomly matched with another participant to determine their payoffs[16]. In the instructions participants were asked to devise their own numerical examples to be sure that they understood the payoff structure. Following this part, each participant received a booklet with randomized instructions and strategy submission sheets for the high and low *R* treatments of the TD along with other unrelated games (not reported here).

For determining their game related payoffs, participants were randomly matched. Participants knew that they will be only paid for two of their decisions. One corresponding to the first (between-subjects) part, and the other one, corresponding to one randomly drawn game from the booklet (within-subjects) part. Feedback on both payoffs was provided once, at the end of the session, in order to minimize the effects of learning and avoid cross-subject dependence. Participants' average earnings in the economic experiment were 15.2 Euros. Instructions are available in Apendix (5.1.2).

[15] May 27, 2009. Each session had the following number of subjects: 28, 29, and 27. The observations of three of these subjects had to be dropped, because of incomplete data in the cognitive ability tasks.

[16] Couples in uneven cohorts were obtained matching one of the subjects twice.

2.4 The dataset

2.4.1 The cognitive tasks measures

As mentioned before, participants completed a total of 5 cognitive tasks. In the following part we will give a short description of the measures obtained in each of the tasks and Table 2.2 summarizes the tasks and their corresponding measures.

- Raven Standard Progressive Matrices Test: We obtained a direct score, corresponding to the number of correct responses (ranging from 0 to 60) additionally using the Spanish scale we classified individuals according to the following population percentiles: 5, 10, 25, 50, 75, 90 and 95.

- Dot Counting Span Task: The total number of accurate responses was used as score (ranging from 0 to 20) as well as the achieved block (0 to 6).

- *N*-Back Task: For each participant, we obtained one performance score (between 0 and 20) for each load: 2 and 3-Back[17].

- Wisconsin Card Sorting Task: We used the percentile of perseverative errors according to the Spanish scale (the lower the amount of perseverative errors, the higher the percentile).

- Stop-Signal Reaction Time Task: We used the stop signal reaction time, which is estimated from the signal response reaction time and the non-signal reaction time.

[17] The performance scores were obtained subtracting the number of false alarms to the number of hits. See the Appendix, for details on the meaning of each measure.

Table 2.2

Summary of cognitive task measures

Task	Var. code	Var. description
Raven Progressive Matrices	RAVEN_SCORE RAVEN_P	Raven direct score Raven score percentile according to Spanish scale
Dot Counting Span	DCS_TOT DCS_BLOCK	Dot Counting total of correct responses Dot Counting block reached
N-Back	2-BACK_SCORE 3-BACK_SCORE	Hits minus false alarms in 2-Back Hits minus false alarms in 3-Back
Wisconsin CST	WIS_PPERSM	Perseverative errors percentile using Spanish scale
Stop-Signal	SS_RT	Stop-Signal response time (miliseconds)

2.4.2 The Traveler's Dilemma decisions

For each participant we have three one shot decisions. The first strategy corresponds to either the low R or the high R version in the TD, depending on his/her randomly assigned treatment, while the second and third strategies correspond to the two versions of the game, in random order. Therefore for the first decision (between-subjects design) we have 40 and 41 independent observations for each treatment (R_L and high R_H, respectively), and 81 dependent observations for each treatment for the second part (within-subjects design). Additionally, since the within-subjects design provided us information about participants' claims across the two treatments we calculated the difference between the two claims (R_L claims minus high R_H claims) as a measure of idiosyncratic *R* parameter sensitivity[18].

2.5 Results

2.5.1 General results

Regarding the cognitive measures (Table 2.3), the correlations within tasks are high, as expected, while correlations across tasks are in most cases not statistically significant. The exception is a mild correlation ($\rho=0.21$, $p<0.10$) between the Dot Counting Span Task reached block (DCS_BLOCK) and the Raven percentile (RAVEN_P). This result

[18] To control for possible ordering effects, we constructed four dummy variables, each one accounting for one specific order in which decisions were taken.

is not surprising, since among the theoretical constructs within theories of information processing, working memory capacity is the one parameter that correlates best with measures of reasoning ability (Oberauer, Schulze, Wilhelm, & Süß, 2005).

Table 2.3

Tasks measures descriptive statistics and spearman rank correlations

		Descriptives		**Correlation Matrix**							
	Var. code	Mean	S.D.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	RAVEN_SCORE	47.2	5.1	1							
2	RAVEN_P	48.7	24.8	**0.95** ***	1						
3	DCS_TOT	13.4	2.1	0.15	0.18	1					
4	DCS_BLOCK	4.3	0.6	**0.19** *	**0.21** *	**0.78** ***	1				
5	2-BACK_SCORE	16.1	3.1	0.00	0.05	0.00	-0.09	1			
6	3-BACK_SCORE	12.7	4.3	0.00	0.00	0.11	0.03	**0.33** ***	1		
7	WIS_PPERSM	45.1	31.3	0.15	0.13	0.15	0.18	0.08	0.04	1	
8	SS_RT	664.0	170.8	-0.02	-0.03	0.10	0.15	-0.15	0.00	-0.07	1

*10% confidence level, **5% confidence level, ***1% confidence level.

Results from the economic experiment, as it can be seen in Figure 2.1, are in line with those obtained by Goeree and Holt (2001); That is, when the *R* parameter, namely 5, individuals' claims tend to concentrate in the higher level (300), while when *R* is high (180) the modal claim is the Nash equilibrium (180)[19]. Nevertheless, our participants' claims are not exclusively clustered into extreme values, specially, in the low reward/penalty treatment, where less than a third chose the maximum claim[20]. We also observe a significant proportion of individuals choosing inner claim values, which are mostly multiples of ten. In fact, for the whole claim range [180-300] concentration in this set of numbers is between 75.6% and 90.7%, which could be explained by prominence theory (Albers, 1999).

[19] The difference between H_R and H_L treatment results is significant at any standard level of confidence using a Kolmogorov-Smirnov test for the between-subjects part, and a Wilcoxon matched-pairs signed-ranks test for the within-subjects part.

[20] 19.5% in the between and 27.4% in the within-subjects part.

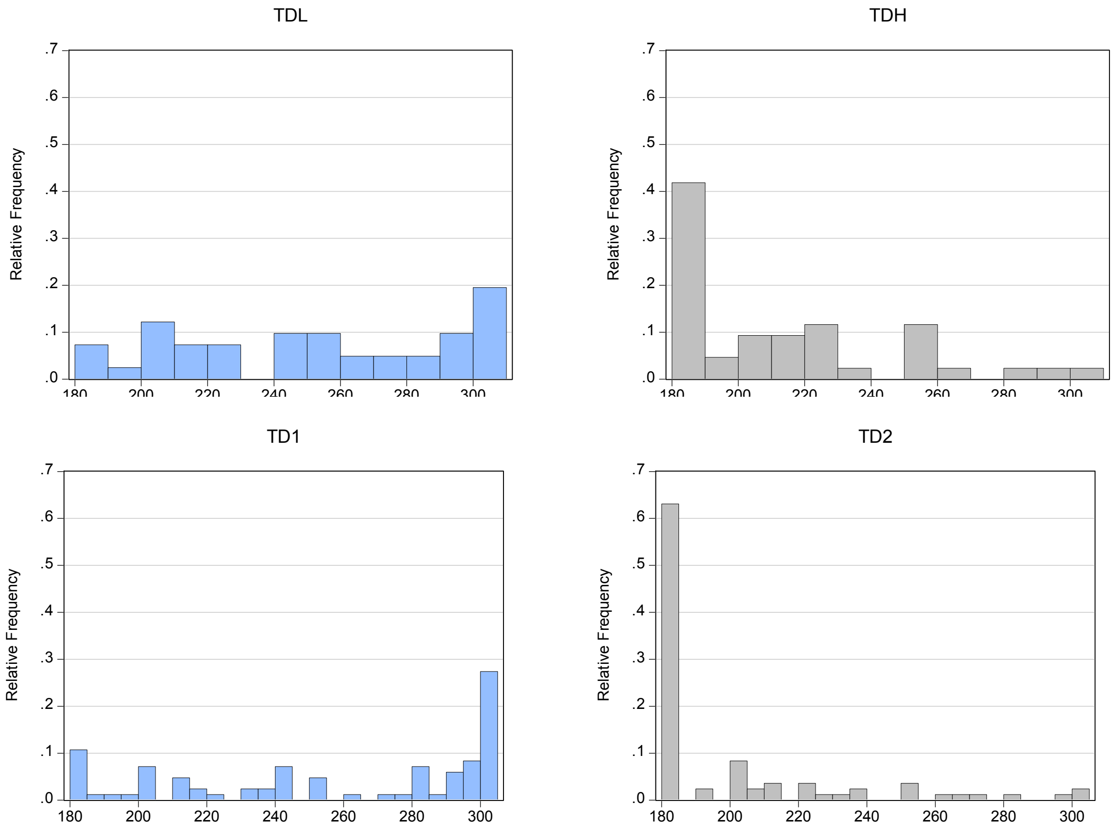

Figure 2.1 *Claims Frequencies in Traveler's Dilemma. Between-Subjects (Top); Within-subjects (Bottom)*

A peak also occurs in the central values [240-250] which can be associated with an embedding bias similar to that reported by Bosch-Domènech an Silvestre (2006) on Holt and Laury (2002) lotteries. Regardless of what causes this dispersion in our data, this property is an advantage, given that we are interested in explaining heterogeneity and not only modal behavior. Finally, comparing the data obtained in the between and within-subjects designs we notice that the distributions are similar across same treatments[21].

To complement the analysis of the within-subjects part, we computed for each participant the difference between their claims in the low and high R treatment, that is, R_L minus R_H. Figure 2.2 displays the frequency distribution of this difference. A positive value captures subjects submitting a higher claim in the R_L treatment than in the

[21] Indeed, using a Two-sample Kolmogorov-Smirnov test we cannot reject claim distributions equality, for both treatments (p>0.10). Also, considering the within-subjects design we ran a Wilcoxon matched-pairs signed-ranks test, for subjects that played the same treatment across parts, and we found that claim medians are not significantly different, for both treatments (p>0.10).

R_H treatment, that is, those who behave according to the intuitive pattern, and vice versa. A difference of zero, on the other hand, corresponds to those players claiming the same amount in both treatments, that is, participants who are insensitive to changes in the *R* parameter.

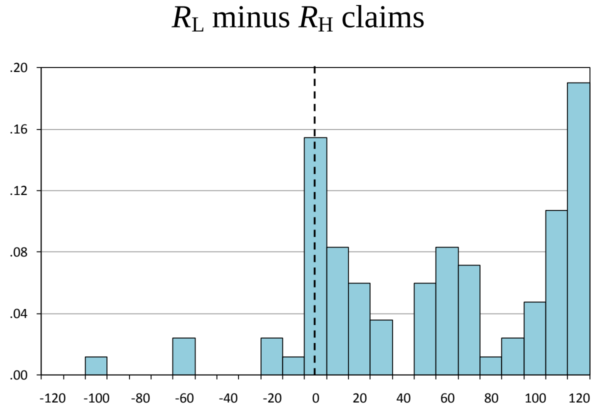

Figure 2.2 *Difference among claims (Within-subjects)*

As Figure 2.2 shows, the great majority of participants (82.2%) submitted higher claims in the R_L treatment with respect to the R_H treatment. The most frequent difference between claims is 120 (19.1%), which corresponds to maximum claim (300) in the R_L treatment and Nash equilibrium claim in the R_H treatment (180). Another peak is observed in the value zero (10.7%) which corresponds to people playing the same strategy in both treatments. It is important to mention, that the majority of this cases (7 out of 9) correspond to players choosing the Nash strategy in both treatments (180, 180). Very few participants (7.1%) played the counterintuitive pattern of higher claims in R_H than in the R_L treatment.

2.5.2 Cognitive abilities and behavior in the TD game

As a first exploratory analysis on the link between cognitive abilities and behavior in the TD game, we computed spearman rank correlations between every task measure and claim levels in each TD game. As shown in Table 2.4, R_L claims are positively correlated with working memory capacity (Dot Counting Span) in both between-subjects and within-subjects part, and with reasoning ability level (Raven) in the within subject part. On the other hand R_H claims are not correlated to any task measure in the between-subjects part, while a negative correlation shows in the within-subjects part, with working memory capacity (2-Back). The most interesting result is that distance between claims in the two *R* treatments is positively correlated with reasoning ability capacity (Raven) and working memory capacity (Dot Counting Span).

Table 2.4

Correlation between task measures and claim levels

Task	Var.	Between-Subject		Within-Subject		
		R_L	R_H	R_L	R_H	$R_L - R_H$
Raven PM	RAVEN_SCORE	0.06	0.05	0.13	-0.09	0.17
	RAVEN_P	0.13	0.04	**0.19** *	-0.07	**0.20** *
Dot Counting Span WM	DCS_TOT	**0.30** *	-0.01	**0.21** *	0.08	**0.21** *
	DSC_BLOCK	**0.30** *	0.16	0.10	0.08	0.11
N-back	2-BACK_SCORE	0.05	-0.04	0.06	**-0.21** *	0.13
	3-BACK_SCORE	-0.03	0.02	0.06	-0.03	0.03
Wisconsin	WIS_PPERSM	0.07	0.11	0.05	0.06	-0.02
Stop-signal	SS_RT	0.03	0.19	0.02	0.12	-0.03
Number of observations		40	41	81	81	81

*10% confidence level.

Therefore it seems that the "contradictory" shift from high to low claims due to a change in the "irrelevant" payoff *R* is not correlated to poor but to high performance in cognitive ability tests. The question then is: is it really that surprising, or "playing reasonably" might be more appealing a concept than "equilibrium play"? To support this conjecture, we matched each possible claim with the actual sample of claims obtained in the experiment. That is, we constructed an ex-post expected payoff curve for each treatment over the strategy space available (Figure 2.3).

Notice that the expected payoff increases with claims in the low *R* treatment and decreases in the high *R* treatment. Therefore, an individual seeking to maximize his payoffs should play high in R_L treatment and vice versa. To be precise, the perfect foresight strategy is to claim 299 when the R=5 and 180 (the Nash equilibrium) when R=180. Moreover, for the latter, the closed claim range 295-299 is more profitable than playing either 294 or 300. Therefore we can define 180 and [295-299] as “strategic claims”, for R_H and R_L, respectively.

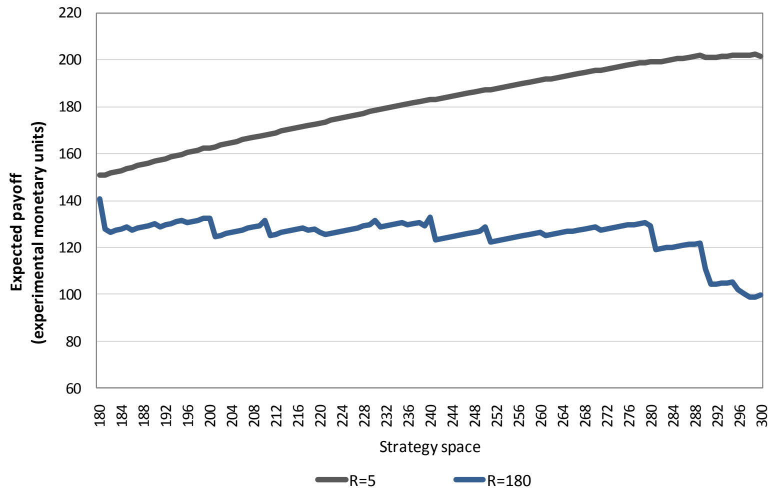

Figure 2.3 *Expected payoffs over the strategy space*

Focusing on participants who played these “strategic claims” and comparing them with the rest, we find the following:

- Scores in the 2-Back Task are higher for participants playing the strategic claims compared to those playing other values, both for R=5 and R=180. (Mann Whitney U test, $p<0.05$ for the within-subjects design).

- Score in the Raven Standard Progressive Matrices of individuals playing the range [295-299] are higher than the rest. To be precise, the Raven score average for them is around 52, which corresponds to the 75th percentile using Raven’s Spanish scale (for ages between 19 and 22), while the average scores for the rest of claims is around 47 (50th percentile). Moreover, median differences, using the Mann Whitney U test are significant ($p<0.05$ and $p<0.01$, for the between and within-subjects parts, respectively).

- We did not find other significant differences in the rest of cognitive abilities tasks scores.

Figures 2.4 and 2.5 show the average value of the cognitive scores for the defined strategic claim values compared to the average for the rest of the claims for those task measures where the difference in scores between the two groups is statically significant.

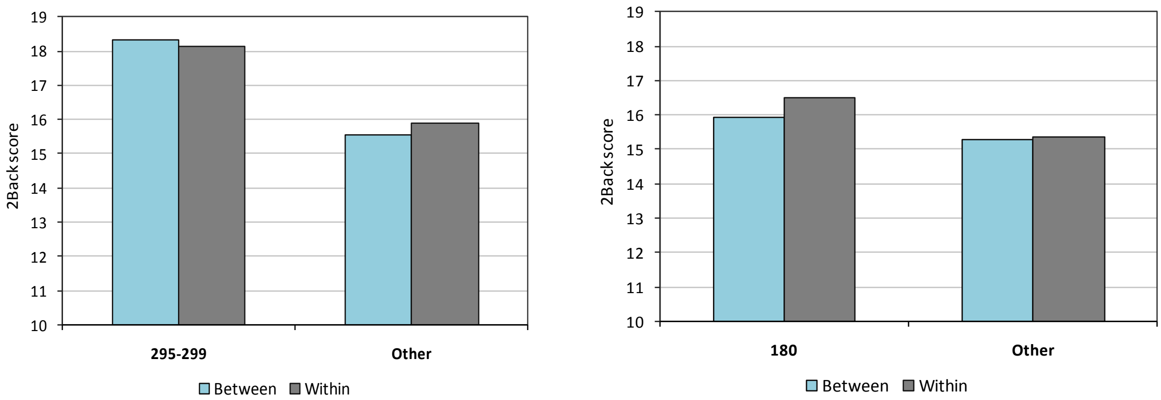

Figure 2.4 *Average 2-Back Task scores by claim range*

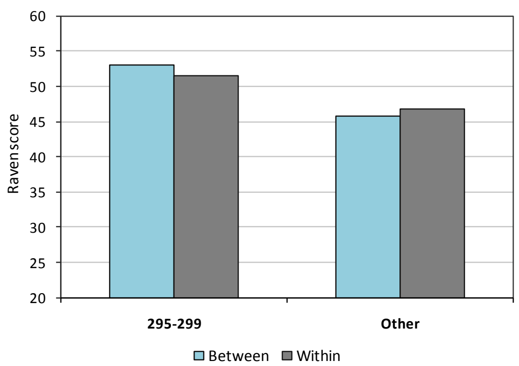

Figure 2.5 *Average Raven scores by claim range*

Therefore results point towards the conclusion that more cognitively sophisticated individuals are indeed more strategically sophisticated players, which does not mean that they play the Nash equilibrium in the low *R* (contradiction) treatment, but rather that they might anticipate modal behavior and best respond to other player actions, hence maximizing their expected payoffs.

To extend the previous analysis to a multivariate framework, that is, to study the joint effect of different cognitive measures on the probability of playing *strategic claims*, we used a Probit specification. The results obtained are summarized in Table 2.5. The dependent variables correspond to dummies that take the value 1, when claims are in the strategic values, that is, [295-299] and 180, for the low *R* and the high *R* treatments, respectively. Due to the sample size we could only include those treatments belonging to the within-subjects design (control variables for ordering were included).

Table 2.5

Probit regression for strategic claims and cognitive abilities

Task	Variable	R_L [295-299]		R_H [180]	
		(1)	(2)	(3)	(4)
Raven PM	RAVEN_P	**0.100** ** (0.048)	**0.051** ** (0.023)	**0.005** (0.007)	**0.004** (0.007)
Dot Counting Span	DCS_TOT	**0.078** (0.252)	**-0.034** (0.176)	**-0.043** (0.092)	**-0.043** (0.093)
N-back	2-BACK2_SCORE	**0.604** ** (0.296)		**0.099** * (0.055)	
	3-BACK_SCORE		**0.082** (0.081)		**0.040** (0.036)
Wisconsin	WIS_PPERSM	**-0.002** (0.016)	**0.006** (0.013)	**-0.004** (0.006)	**-0.004** (0.006)
Stop-Signal	SS_RT	**-0.014** (0.016)	**-0.012** (0.011)	**-0.003** (0.004)	**-0.003** (0.004)
	C	**25.425** (17.830)	**19.518** (13.035)	**-6.522** (4.635)	**-4.452** (4.283)
McFadden R-squared		**0.602**	**0.425**	**0.226**	**0.206**

Dependant variable: Dummy that takes value=1 when the claim corresponds to the indicated range and zero otherwise.
*10% confidence level, **5% confidence level, ***1% confidence level.
Standard errors in parentheses.

As our previous findings suggested, for the low R treatment, a higher Raven score implies a higher probability of using the minimal undercutting strategy. Additionally, we find that a higher working memory capacity, measured by 2-Back Task score implies a higher probability of playing strategically in both the low and the high R treatments (Models 1 and 3).

2.6 Conclusion

Experimental evidence in controlled strategic interaction situations has inspired "behavioral" game theory models in which a better description of observed results is achieved by relaxing one or more of the main assumptions of classic game theory (self-interest, perfect rationality and common knowledge-belief in others rationality). For competitive dominance solvable games, such as the TD, most of the behavioral game theory models explicitly address the issue of cognitive limitations as one of the reasons for non-equilibrium behavior.

Surprisingly scarce attempts have been made to obtain external measures of potentially relevant cognitive traits in order to test whether they correlate with observed behavior in strategic interaction contexts. Notable exceptions are the works by Devetag and Warlien (2003) and Rydval et al. (2009) who investigate behavior in different simple dominance solvable games, and Burnham et al. (2009) and Brañas et al. (2011) for the beauty contest game. The main two conclusions of these studies are that participants with better performance in different cognitive ability tests play more accordingly to standard game theory prescriptions and, that reasoning errors as well as dominance violations are associated with lower levels of cognitive abilities.

In this essay we have studied behavior in the Traveler's Dilemma game and applied a parallel set of cognitive tasks. Unlike the games previously studied, the Traveler's Dilemma, as well as other dominance solvable games, such as the prisoners dilemma and the centipedes game, has a structure that makes iterated reasoning decrease payoffs, therefore the Nash equilibrium strategy is sub-optimal compared to the cooperative outcome. Given this, should one expect individuals with higher cognitive abilities to behave closer to standard game theoretic prescription in such type of games?

The result obtained in this essay reveal that more cognitively sophisticated individuals are indeed more strategically sophisticated players, which does not mean that they play more accordingly to the standard game theory prediction when the reward/penalty parameter is low, but that they might anticipate modal behavior and best respond to other player actions, using the minimal undercutting strategy and end up achieving the largest payoffs. Also we find evidence that players with higher cognitive abilities are more likely to exhibit the contradictory pattern of switching across *treasure* and *contradiction* treatment. This evidence goes against the conjecture that this switching is due to poor understanding of the strategic nature of the game, suggesting just the contrary: The unexpected switch is the result of a cognitively demanding forecast of what others will do under each parameter, as corroborated by an exercise of fictitious matching of observed individual behavior with that of our participants' population.

Our results are in line with Bayer and Renou (2011) who found that those subjects that were able to perform more steps of iterated reasoning were more strategically sophisticated and outperformed Nash equilibrium payoffs in several dominance solvable games. And also in some measure with Jones (2008) who from a meta-study of repeated prisoner's dilemma experiments, reports that higher levels of cooperation are achieved in universities that admit students with higher SAT scores.

2.7 References

Albers, W. (1999). Prominence Theory as a Tool to Model Boundedly Rational Decisions. In G. Gigerenzer & R. Selten (Eds.), *Bounded Rationality, The Adaptive Toolbox* (pp. 297-318). Cambridge: MIT Press.

Anderson, S. P., Goeree, J. K., & Holt, C. A. (2002). The logit equilibrium: A perspective on intuitive behavioral anomalies. *Southern Economic Journal*, *69*(1), 21–47.

Baddeley, A. (1992). Working Memory. *Science*, *255*(5044), 55.

Basu, K. (1994). The Traveler's Dilemma: Paradoxes of Rationality in Game Theory. *American Economic Review*, *84*(2), 391–395

Basu, K., Becchetti, L., & Stanca, L. (2011). Experiments with the Traveler's Dilemma: welfare, strategic choice and implicit collusion. *Social Choice and Welfare*, *37*(4), 575-595.

Bayer, R. C., & Renou, L. (2011). Cognitive abilities and behavior in strategic-form games. *Journal of Economic Behavior & Organization*.

Becker, T., Carter, M., & Naeve, J. (2005). *Experts playing the Traveler's Dilemma* (Discusion paper 252/2005). Stuttgart, Germany: Department of Economics, Universität Hohenheim. Retrieved from: https://www.uni-hohenheim.de/RePEc/hoh/papers/252.pdf

Berg, E. A. (1948). A Simple Objective Technique for Measuring Flexibility in Thinking. *Journal of General Psychology*, *39*, 15-22.

Bosch-Domènech, A., & Silvestre, J. (2006). *Risk aversion and embedding bias* (Economics Working Papers 934). Barcelona, Spain: Department of Economics and Business, Universitat Pompeu Fabra. Retrieved from: http://www.econ.upf.edu/docs/papers/downloads/934.pdf

Brañas-Garza, P., Espinosa, M. P., & Rey-Biel, P. (2011). Travelers' types. *Journal of Economic Behavior & Organization*, *78*(1-2), 25-36.

Brañas-Garza, P., García-Muñoz, T., & Hernan, R. (2011). *Cognitive effort in the Beauty Contest Game* (Working papers 11-08). Orange, CA: Chapman University Economic, Science Institute. Retrieved from: http://www.chapman.edu/ESI/wp/Hernan-CognitiveEffortBeautyContestGame.pdf

Burnham, T. C., Cesarini, D., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Higher cognitive ability is associated with lower entries in a p-beauty contest. *Journal of Economic Behavior & Organization*, *72*(1), 171-175.

Cabrera, S., Capra, C. M., & Gómez, R. (2007). Behavior in one-shot traveler's dilemma games: model and experiments with advice. *Spanish Economic Review*, *9*(2), 129-152.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A Cognitive Hierarchy Model of Games. *Quarterly Journal of Economics*, *119*(3), 861-898.

Capra, C. M., Goeree, J. K., Gomez, R., & Holt, C. A. (1999). Anomalous behavior in a traveler's dilemma? *American Economic Review*, *89*(3), 678-690.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, *97*(3), 404.

Colman, A. M. (2003). Depth of strategic reasoning in games. *Trends in Cognitive Sciences*, *7*(1), 2-4.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547-552.

Costa-Gomes, M., & Crawford, V. P. (2006). Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. *American Economic Review*, *96*(5), 1737-1768.

Dempster, F. N. (1991). Inhibitory Processes: A Neglected Dimension of Intelligence. *Intelligence*, *15*, 157-173.

Devetag, G., & Warglien, M. (2003). Games and phone numbers: Do short-term memory bounds affect strategic behavior? *Journal of Economic Psychology*, *24*(2), 189-202.

Eichberger, J., & Kelsey, D. (2011). Are the treasures of game theory ambiguous? *Economic Theory*, *48*(2-3), 313-339.

Eling, P., Derckx, K., & Maes, R. (2008). On the historical and conceptual background of the Wisconsin Card Sorting Test. *Brain and cognition*, *67*(3), 247-53.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology*, *128*(3), 309-331.

Erlei, M. (2008). Heterogeneous social preferences. *Journal of Economic Behavior & Organization*, *65*(3-4), 436-457.

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

Goeree, J. K., & Holt, C. A. (2001). Ten Little Treasures of Game Theory and Ten Intuitive Contradictions. *American Economic Review*, *91*(5), 1402-1422.

Goeree, J. K., & Holt, C. A. (2004). A model of noisy introspection. *Games and Economic Behavior*, *46*(2), 365-382.

Halpern, J. Y., & Pass, R. (2012). Iterated regret minimization: A new solution concept. *Games and Economic Behavior*, *74*(1), 184-207.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sort Test Manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources, Inc.

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

Jones, G. (2008). Are smarter groups more cooperative? Evidence from prisoner's dilemma experiments, 1959–2003. *Journal of Economic Behavior & Organization*, *68*(3-4), 489-497.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352-358.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning Ability is (Little More Than) Working-Memory Capacity?! *Intelligence*, *14*(4), 389-433.

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, *10*, 6-38.

Nagel, R. (1995). Unraveling in Guessing Games: An Experimental Study. *American Economic Review*, *85*(5), 1313-1326.

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H. M. (2005). Working Memory and Intelligence--Their Correlation and Their Relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychonomic Bulletin*, *131*(1), 61-65.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*, 46-59.

Palacios-Huerta, I., & Volij, O. (2009). Field Centipedes. *American Economic Review*, *99*(4), 1619-1635.

Pickering, S. J., Baqués, J., & Gathercole, S. E. (1999). *Bateria de tests de memòria de Treball* (mimeo). Barcelona, Spain: Laboratori de Memòria del U.A.B.

Raven, J. C. (1960). *Guide to using the Standard Progressive Matrices.* London: Lewis.

Rigas, G., Carling, E., & Brehmer, B. (2002). Reliability and validity of performance measures in microworlds. *Intelligence, 30*(5), 463-480.

Rubinstein, A. (2006). Dilemmas of an Economic Theorist. *Econometrica*, *74*(4), 865-883.

Rubinstein, A. (2007). Instinctive and cognitive reasoning: Response times study. *The Economic Journal*, *117*(523), 1243-1259.

Rydval, O., Ortmann, A., & Ostatnicky, M. (2009). Three very simple games and what it takes to solve them. *Journal of Economic Behavior & Organization*, *72*(1), 589-601.

Stahl, D. O., & Wilson, P. W. (1995). On Players' Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior*, *10*, 218-254.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127-154.

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in cognitive sciences*, *12*(11), 418-24.

Essay 2

3. On the Cognitive Foundations of Mixed Strategy Equilibrium Failure in the Lab[22]

3.1 Introduction

Economists have thoroughly studied situations in which the predictions of game theory are challenged by the behavior of real agents. In the vast majority of the existing studies a game with one or more pure strategy equilibria has been used as the test bed in the laboratory. However, there are a number of games in which there is no pure strategy equilibrium. In these cases, the traditional game theory equilibrium solution involves the use of mixed strategies, that is, players randomizing over actions (pure strategies) in a way which cannot be predicted and therefore exploited by the opponent (Nash, 1951; von Neumann & Morgenstern, 1944). To be precise, in a mixed strategy Nash equilibrium players select randomly among available actions from a probability distribution that makes the opponent(s) indifferent between any of the actions available to them. A difficulty with testing this concept with real data is that the existence of equilibrium strategy distributions cannot be verified unless we make one of the two following strong assumptions: (1) either that a sequence of strategies produced by the same subject is randomly extracted from equilibrium strategy distributions and are thus serially uncorrelated or (2) that the choices by a population of subjects follows the distribution predicted for each one of them.

In this essay, we adopt the second assumption by implementing a one-shot asymmetric matching pennies game, a two-by-two game in which one player wins if decisions match and the other wins if decisions differ. The game has a unique equilibrium in mixed strategies, but experimental evidence shows that when large payoff-asymmetries are present, equilibrium predictions are systematically violated both in one-shot and repeated implementations (Goeree & Holt, 2001; Goeree, Holt, & Palfrey, 2003; McKelvey, Palfrey, & Weber, 2000; Ochs, 1995). Understanding and explaining the sources of the systematic departures from mixed strategy equilibrium in this type of

[22] This essay is the result of join work with M. Teresa Bajo-Molina, Pablo Brañas-Garza, Nikolaos Georgantzís, and Julia Morales. Financial support by the Junta de Andalucía (P07-SEJ-03155) and the Spanish Ministry of Science and Innovation (ECO2008-04636/ECON; ECO 2010-17049) is gratefully acknowledged.

framework is relevant since many economic situations, such as auditing, can be modeled as asymmetric non-zero sum games. Moreover, these strategic situations often occur without clear precedents, therefore equilibrium (initially) depends on strategic thinking, which in turn depends on subject's cognitive abilities, among other factors. Following this line of thought, our approach aims at testing the role of specific idiosyncratic factors and cognitive abilities as an explanatory factor of deviations from the predicted mixed equilibrium behavior in an asymmetric matching pennies game. More specifically, our objective is to address whether or not there are idiosyncratic cognitive motivators which can explain the probability that a subject's one-shot choice is on one or the other side of the distribution.

Implicitly, our testable hypotheses stem from the juxtaposition of the two main alternative explanations for initial response deviations from the standard mixed strategy equilibrium (iterated noisy introspection and level-k reasoning). In this framework, errors could be expected to relate with low performance in cognitive tests while higher level-k reasoning with high performance in cognitive tests. Our findings show that psychological factors, like a subject's lack of inhibition and working memory may influence a subject's choices in a way that induces systematic deviations from the predicted mixed strategy equilibrium distributions.

The rest of the essay is organized as follows. Section 3.2 presents a literature review on theoretical interpretations and the empirical tests of the mixed strategy equilibrium. Section 3.3 describes methodological aspects of the experiment. Section 3.4 presents results. And Section 3.5 presents the concluding remarks.

3.2 Theoretical and empirical background

Although "the theory of mixed-strategy play, including von Neumann's Minimax Theorem and the more general notion of a Nash equilibrium in mixed strategies, remains the cornerstone of our theoretical understanding of strategic situations that require unpredictability" (Walker & Wooders, 2001, p. 1521), mixed-strategy play as shown to be problematic both in theory and practice.

As Rubinstein (1991, p. 912) noted "The concept of mixed strategy has often come under heavy fire". According to Harsanyi (1973, p.1) "Equilibrium points in mixed strategies seem to be unstable, because any player can deviate without penalty from his equilibrium strategy even if he expects all other players to stick to their.” Aumann (1987, pp. 15-16) noted that “mixed strategy equilibria appear quite special and rather unnatural. They imply that the players always act as if they all had the same beliefs about what all other players will do, and as if these beliefs were common knowledge… In particular, they act as if each player *i* always knew exactly what each other player believes about his (i's) actions.”

To overcome these difficulties, different interpretations of the mixed strategy Nash equilibrium (henceforth MSNE) have been given, nevertheless a general consensus is still missing (Gallice, 2007). Harsanyi (1973) proposed the purification theorem, according to which mixed strategy equilibria are explained as being the limit of pure strategy equilibria for a game with small random fluctuations in the payoffs and in which the payoffs of each player are known only by themselves. The idea is that the predicted MSNE of the original game emerges because the random fluctuations in payoffs will make players use their pure strategies approximately with the probabilities prescribed by the MSNE. Aumann (1987) goes further, claiming that even in the absence of private information, players are still uncertain about the opponents’ moves. He proposes the idea of a *correlated equilibrium*, where each player chooses a pure strategy with no intention to randomize. The probability distribution that the MSNE assigns to each player can be directly interpreted as the uncertainty in the beliefs of other players about his choice. Another interpretation of the MSNE is that the frequencies of pure strategies represent the steady state probabilities when games are played in large populations. Therefore, although each player is choosing a pure strategy in the entire population the MSNE distribution emerges (Gallice, 2007).

On the empirical side, the interest in using experimental methods to test the predictive power of the mixed strategy equilibrium was revived by O’Neill’s (1987) PNAS paper (Rosenthal, Shachat, & Walker, 2003). Most empirical studies after O’Neill’s until nowadays are focused on testing the strictest definitions of MSNE by analyzing repeated interaction in games with unique MSNE and looking for evidence in favor or against individual randomization and the predicted distribution of choice frequencies.

Even for very simple zero-sum games that are played repeatedly the empirical evidence from laboratory experiments and natural experiments is inconclusive. In general terms, most of laboratory experiments with undergraduate students provide weak support for the mixed strategy play, particularly at an individual level (Brown & Rosenthal, 1990; O'Neill, 1987; Rapoport & Boebel, 1992; Shachat, 2002). Long run aggregate frequencies of play are usually close to equilibrium predictions, but most studies find evidence against player randomization of choices. Usually, individuals tend to over-switch among options (Brown & Rosenthal, Rapoport & Boebel, 1992; Mookherjee & Sopher, 1994) and very infrequently under-switching is observed (Rosenthal et al., 2003).

Another empirical approach to test mixed-strategy play has been to use data from sports where competitive zero-sum strategic situations occur as part of the game. Papers using this approach analyze situations such as serve and return play in Tennis (Hsu, Huang, & Tang, 2007; Walker & Wooders, 2001), penalty shots in Soccer (Azar & Bar-Eli, 2011; Chiappori, Levitt, & Groseclose, 2002; Palacios-Huerta, 2003) and pitch type in Baseball (Kovash & Levitt, 2009). The alleged advantage to use this approach is that expert players should conform best to theoretical predictions. In fact, it seems to be the case that the observed frequencies of play are more consistent with the mixed-strategy predictions than in the laboratory experiments. On the other hand, evidence on the randomness of play is supported only by some of the studies (Palacios-Huerta, 2003) while several others find strong evidence of negative serial autocorrelation (Kovash & Levitt, 2009; Walker & Wooders, 2001).

The question whether or not an experts' ability to mix their strategies transfers from the playing field to the laboratory remains an open question. Palacios-Huerta (2003) find that Spanish professional soccer players conform to the theory in the lab when playing two abstract games, but Wooders (2010) reexamines the data and reports that the play of professionals is inconsistent with the minimax hypothesis in several respects.

While the empirical evidence for symmetric zero-sum games is rather inconclusive, in experiments involving non zero-sum games with considerable payoff asymmetries equilibrium play predictions are systematically rejected, even at an aggregate level. Ochs (1995) investigates behavior in three matching pennies games. One symmetric (zero-sum) and the other two with a payoff asymmetry (non-zero-sum). He reports that in the asymmetric games players select their salient own-payoff option considerably more than one-half of the time (the Nash prediction). Ochs (1995) results are replicated by McKelvey, Palfrey, and Weber (2000) and explained using Quantal Response Equilibrium. Similarly, Goeree, Holt, and Palfrey (2003) report results for a ten-period repeated matching pennies games. They find results which are qualitatively similar but less dramatic than those found by Goeree and Holt (2001) who found strong "own-payoff" effects that are not predicted by the Nash equilibrium (Goeree & Holt, 2001).

The results from these studies show that in MPG with an unbalanced payoff structure the mixed strategy predictions are violated systematically with players' choices responding to their own salient payoffs in an intuitive manner. According to Goeree and Holt (2004) this happens because Nash predictions do not pick up systematic "own-payoff" effects that alter quantitative but not qualitative payoff comparisons. The effect of payoff magnitudes on equilibrium behavior in games was first modeled in a general way by McKelvey and Palfrey (1995) who proposed the Quantal Response Equilibrium (QRE). According to this theory, there are two effects of increased payoff magnitude in games: a) a direct effect: increasing the magnitude of incentives will reduce decision errors by subjects; b) an indirect general equilibrium effect: decision errors of one subject will affect the payoffs of the other subject, this in turn will affect the choice frequencies of that subject which then feedback on the original subject. An equilibrium with these decision errors is a QRE. Thus, changes in payoff magnitudes can change the QRE of the game, even if there is no change in the standard (Nash) equilibrium of the game (McKelvey et al., 2000).

Using the same approach Anderson, Goeree, and Holt (2002) incorporate the payoff-asymmetry effects by introducing noisy behavior via a logit probabilistic choice function. In the resulting *logit equilibrium*, behavior is characterized by a probability distribution that satisfies a "rational expectations" consistency condition. The beliefs that determine the players' expected payoffs match the decision distributions that arise from applying the logit rule to those expected payoffs. For one-shot games (Goeree & Holt, 2004) propose the *iterated noisy introspection* were they relax the equilibrium condition of consistency of actions and beliefs by introducing a process of iterated conjectures. And, for a repeated version of the Asymmetric MPG Goeree et al., (2003) propose a model which incorporates risk aversion into a QRE.

Other models have been used to explain deviations from MSNE in asymmetric MPG. Among them we can mention the Cognitive Hierarchy (CH) Model (Camerer et al., 2004) where players engage in discrete steps of reasoning. Each actor belongs to a hierarchy (depending on the number of steps of reasoning he uses) and has accurate beliefs about the relative frequencies of those below their hierarchy. For one-shot games (Camerer, Ho, & Chong, 2002), mutual consistency is relaxed by assuming that players use k steps of reasoning with frequency f(k), where f is a one-parameter Poisson distribution. Camerer, Ho, and Chong (2002) report that the CH model prediction for an Asymmetric MPG is very close to experimental data[23]. On the other hand, Boylan and Grant (2008) consider fairness-based payoff transformations. They show, however, that such preferences do not predict the observed behavior in the experimental data of Goeree and Holt (2001), which in turn can be better explained using QRE (Eichberger & Kelsey, 2011).

Although all those models improve MSNE predictions for games with unbalanced payoffs, they infer their explanations from the data itself, not explicitly obtaining their explanatory variables through subjects' behavior in an appropriately designed external task. Thus they are showing that the explanation is sufficient but not testing it as such.

3.3 Method

3.3.1 The Asymmetric Matching Pennies Game

As in a usual Matching Pennies Game (MPG), in the asymmetric version of the game the motivations for the two players are exactly opposite. In order not to be exploited by the opponent, neither player should favor one of their strategies. Given this, there is no equilibrium in pure strategies. The game has a unique equilibrium in mixed strategies in which each player randomizes over her two alternatives.

As Figure 3.1 shows, we used the same payoff structure that was reported in Goeree and Holt's (2001) treasures and contradictions paper. In this version of an Asymmetric MPG, the Row Player has a high payoff if the combination *Top-Left* is played, while

[23]The Level-k models (Costa-Gomes, Crawford, & Broseta, 2001) also explain deviations from MSNE in Asymmetric MPG games, but since they assume that each players thinks everyone is one step below them, the model predictions are insensitive to the amount of the asymmetry, while behavioral data shows that the frequency of choice of the salient own-payoff alternative is positively correlated with the size of the asymmetry. The Level-k models are, on the other hand, well suited to explain deviations that arise as a consequence of non-neutral framing of locations (Crawford & Iriberri, 2007) in games similar to the Matching Pennies (Hide-and-Seek Games).

Column Player has the usual symmetric payoff across combinations. The game has no equilibrium in pure strategies.

	Left	Right
Top	320 , 40	40 , 80
Bottom	40 , 80	80 , 40

Figure 3.1 *The payoff matrix for the asymmetric MPG.*

Row's expected payoff for *Top* (U_T) and *Bottom* (U_B) are a function of Column's probability of choosing *Right* (p_R) and *Left* ($1- p_R$). Therefore Row's expected payoffs are:

(1) $$U_T = 320(1- p_R) + 40(p_R) = 320 - 280p_R$$

(2) $$U_B = 40(1- p_R) + 80(p_R) = 40 + 40p_R$$

The difference in these expected payoffs is:

(3) $$U_T - U_B = (320 - 280p_R) - (40 + 40p_R) = 280 - 320p_R.$$

Row is indifferent if the expected payoff difference is zero, i.e. if $p_R = 7/8$.

On the other hand, Column expected payoffs for *Left* (U_L) and *Right* (U_R) are a function of Row's probability of choosing *Bottom* (p_B) and *Right* ($1- p_B$):

(4) $$U_L = 40(1- p_B) + 80(p_B) = 40 + 40p_B$$

(5) $$U_R = 80(1- p_B) + 40(p_B) = 80 - 40p_B$$

The difference in these expected payoffs is:

(6) $$U_L - U_R = (40 + 40p_B) - (80 - 40p_B) = 80p_B - 40.$$

Column is indifferent if the expected payoff difference is zero, i.e. if $p_B = 1/2$.

As it can be seen in the best response function diagram, Figure 3.2, Row's optimal decision is to choose *Top* if Column probability of choosing *Right* is less than 7/8. While Column's optimal decision is to choose *Right* if Row probability of choosing *Bottom* is less than 1/2.

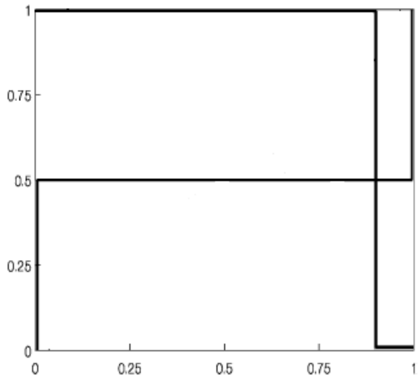

Figure 3.2 *Best response function diagram for our asymmetric matching pennies game*

The intersection of these two curves, at p_R= 7/8 and p_B = 1/2, is the (unique) MSNE of this game. Therefore, the Column Player should randomize between *Right* and *Left* with probabilities 7/8 and 1/8, respectively, while the Row Player should randomize over each decision with equal probability.

3.3.2 Implementation

3.3.2.1 *The game*

We used a within-subjects design, that is, participants played each role once (Row and Column) in random order. Both decisions are considered one-shot, since feedback was not provided in-between, and participants knew they will be randomly matched with different pairs in each game. The game was not presented in matrix form (see instructions in the Appendix 5.2.2) and the choices were labeled as A (*Top* and *Left*) and B (*Bottom* and *Right*) following the original instructions of Goeree and Holt (2001) and Ochs (1995).

3.3.2.1 *The cognitive tasks*

We used a battery of five tasks to measure four different aspects of cognition: Non-verbal intelligence, working memory, flexibility of thinking and response inhibition.

We selected one measure of general non-verbal intelligence, the Raven Standard Progressive Matrices test by Raven (1960), because it can capture the ability to deal with novelty and adapt thinking to a new cognitive problem. And according to psychological research this measure is expected to correlate highly with performance in complex task (Gonzalez, Thomas, & Vanyukov, 2005). Therefore it is a natural candidate to account for subjects' reasoning and general understanding of the game.

In this test, participants were presented with sixty consecutive items, each one consisting of a matrix containing black-and-white abstract figures with one missing figure. For each matrix, response choices (figures) were presented. The participant's task was to choose among the response figures the one that best completed the pattern. The sixty items were divided into five twelve items sets (A to E), with items within a set becoming increasingly difficult. Participants had as much time as they needed to complete this task. The total number of correctly answered items was used as the Raven's score.

Secondly, we implemented two tasks that measure different aspects of working memory. Working memory (henceforth WM) is defined by Baddeley (1992) as a brain system that provides temporary storage and manipulation of the information necessary for complex cognitive tasks. WM is involved in the selection, initiation, and termination of information-processing functions such as encoding, storing and retrieving data. Limitations in WM capacity are an inherent characteristic of humans and therefore also a limitation in the strategic reasoning process.

The first WM task we implemented is a counting span task (Pickering et al., 1999) which simultaneously demands to store and manipulate information. For this task a display booklet was placed in front of each participant consisting of pages which were each showing an area that contained either three, four, five or six dots. The participant was instructed to count aloud the number of dots and remember the count total for later recall. When the participant finished counting, the experimenter presented the next display. After a number of displays had been presented, the experimenter asked the participant to recall all the dot count totals in the order in which the displays were presented. The number of displays per series ranged from two to six, and they were presented in increasing order. Four series of each length were performed for a total of 20 series. The total number of correct responses was used as the dot-counting score (ranging from 0 to 20).

The *N*-Back task (Kirchner, 1958), on the other hand, is a more complex WM task since according to Owen, McMillan, Laird, and Bullmore (2005) it "requires on-line monitoring, updating, and manipulation of remembered information and is therefore assumed to place great demands on a number of key processes within working memory" (p. 47). We implemented a computerized version of Kirchner (1958) using E-prime (Schneider, Eschmann & Zuccolotto, 2002). In this task, each participant was presented with a sequence of stimuli (phonologically distinct letters from the alphabet), and the task consisted of pressing a YES key when the letter in the screen matched the one from *N* steps earlier in the sequence and a NO key when it did not match the one from *N* steps earlier in the sequence. We presented three memory loads *N*, this is, 1-Back, 2-Back and 3-Back. For each memory load n the participant performed a practice block (non scored sequence of 20 letters) to get familiar with the task, and two critical blocks (scored sequences of 30 letters each) in which there were 10 target and 20 no-target stimuli (20 target and 40 no-target stimuli per memory load). Individuals' scores obtained for each load consisted of the number of correct target responses (hits), incorrect target responses (misses), correct non target responses (right rejections), incorrect non target responses (false alarms) and missing responses, as well as reaction times for target, no-target and correct responses.

Another aspect of cognition that we assessed is flexibility of thinking, which according to Cañas, Quesada, Antolí, and Fajardo (2003) is the human ability to adapt cognitive processing strategies to face new and unexpected environmental conditions. For assessing this cognitive dimension, we used Heaton, Chelune, Talley, Kay and Curtiss (1993)'s implementation of Grant and Berg (1948)'s Wisconsin Card Sorting Task. In this task the experimenter presented four stimulus cards to the participant. The attributes of the cards were different in color (red, green, blue, or yellow), number (1, 2, 3 or 4) and shape (circle, cross, star or square). After observing the four cards, the participant was given a stack of additional cards and was then requested to put each card under one of four stimulus cards and to deduce the matching principle on the basis of feedback (correct, incorrect). We used twice each possible matching principle, with the following order: color, shape, number, color, shape, number. Each matching principle stayed the same until the participant correctly performed 10 consecutive matches, at which point the matching principle was changed (e.g., to shape). The task began and continued until either the participant had successfully achieved the 6 matching criteria or until the total number of target cards reached 128. The main dependent measure was the number of classical perseverative errors, which was the number of times participants failed to change matching criterion when the category changed and kept sorting the cards according to the previous, no longer correct matching principle.

Finally, we assessed an important executive control function, response inhibition or the ability to suppress a pre-potent response. We considered important to include this executive control function, since response inhibition allows appropriate responses to meet complicated task demands and adaptation to changing environments. The measure we used to asses inhibitory control was the Stop-Signal task.

In this task subjects have to refrain from responding when a secondary stimulus is presented. We applied Verbruggen and Logan (2008) STOP-IT computerized task. In this task, participants were presented with a series of visual stimuli (squares or circles) and occasionally they heard a tone (stop-signal). Each participant had to perform a primary visual reaction time task, which was to press a key when a circle appeared in the screen and another when the square appeared; while occasionally a tone indicated them to stop their response to the primary task. Participants performed 3 blocks, in each block a sequence of 64 visual stimuli was presented, half of them circles and the other half squares, and a quarter of the times the visual stimulus was accompanied by a phonetic stop-signal after a variable stop-signal delay. On both no-stop-signal trials and stop-signal trials, the stimulus remained on the screen until participants responded or until the maximal RT had elapsed. Participants were instructed to respond as quickly and accurately as possible to the go stimulus on no-stop-signal trials. The instructions emphasized that they should not slow down to wait for possible stop signals, since the software would detect this behavior and delay the appearance of the signal. The main result was whether or not participants withhold their response to the primary task when the stop signal occurs. This is measured by the stop signal reaction time, which is estimated from the signal response reaction time and the non-signal reaction time. Table 3.1 summarizes the information provided in this section.

Table 3.1

Summary of cognitive abilities assessed and corresponding tasks measures

Cognitive ability	Task	Var. code	Var. description
Non-verbal reasoning	Raven Progressive Matrices	RAVEN_SCORE RAVEN_P	Raven direct score Raven percentile according to Spanish scale
Working memory	Dot Counting Span	DCS_TOT DCS_BLOCK	Dot Counting total of correct responses Dot Counting block reached
	N-Back	2-BACK_SCORE 3-BACK_SCORE	Hits minus false alarms in 2-Back Hits minus false alarms in 3-Back
Flexibility of thinking	Wisconsin CS	WIS_PPERSM	Perseverative errors percentile using Spanish scale
Response inhibition	Stop-Signal	SS_RT	Stop-Signal response time (miliseconds)

3.3.3 Participants and procedures

A total of 81 subjects participated in our study. All of them were first and second year undergraduate economics and business students from the University of Granada[24]; (53 females and 28 males) aged on average 19.18 years (S.D.=1.55, age range=18-27).

Three consecutive experimental sessions were conducted in the Economics and Business Faculty. Sessions took place the same day[25], in order to minimize the possibility that participants in different experimental sessions might share relevant information that could affect their decisions.

After submitting strategies for a first unrelated game (Traveler's Dilemma) each participant received a booklet with randomized instructions and strategy submission sheets for both roles of the Asymmetric MPG along with other games (not reported here).

For determining their game related payoffs, the participants were randomly matched. The participants knew that they will only be paid for two of their decisions. One from the first part and another corresponding to one randomly drawn game from the booklet. In order to minimize the effects of learning and avoid cross-subject dependence, feedback was provided at the end of the session. The experiment was administered in a paper-and-pencil format and lasted around 50 minutes. The participants' average

[24] To the best of our knowledge none of them had any prior formal training in game theory.

[25] May 27, 2009. Each session had the following number of subjects: 28, 29, and 27. The observations of three of these subjects had to be dropped, because of incomplete data in the cognitive ability tasks.

earnings in the economic experiment were 15.2 Euros. The instructions are available in the Appendix (5.2.2).

The cognitive tasks were administered at the University of Granada Experimental Psychology Lab. For completing them the participants assisted to two sessions[26], which lasted around 45 minutes each. The administration of the tasks to the total of participants was done during a period of five weeks. The order in which the tasks were completed by the participants in each session was counterbalanced so that each task was run in each position (first, second, third, fourth and fifth) equally often. The participants completed the computerized tasks (Raven Standard Progressive Matrices, *N*-Back and Stop-Signal) individually in rooms with one or two participants, monitored by one experimenter, while tasks that required to be directly administered by the experimenter (Dot Counting Span and Wisconsin Card Sorting) were performed in a separate room. Each participant received a flat fee of 10 Euros for his/her participation in this part.

For all participants we have their scores in the cognitive tasks and the two choices in the Asymmetric MPG, as well as information about their age and gender.

3.4 Results

3.4.1 General results

Our results from the main experiment are similar to those reported by Goeree and Holt (2001) with respect to Row Players, that is, a high proportion of individuals (88.9%) selected the high own-payoff option *Top*, diverging from the mixed strategy Nash equilibrium prediction (50.0%). Nevertheless, when playing as Column Player, surprisingly only 42.0% choose the option *Right*, not anticipating Row's bias towards the option *Top*, contrary to what is observed in Goeree and Holt (2001) where 84.0% choose this option, neither close to the mixed strategy Nash equilibrium prediction of 87.5% (7/8).

[26] The two sessions took place in consecutive days.

	Left: 58.0%	Right: 42.0%
Top: 88.9%	320 , 40	40 , 80
Bottom: 11.1%	40 , 80	80 , 40

Figure 3.3 *Observed choice percentages for each strategy.*

Regarding expected payoffs for the observed choice frequencies (Table 3.2), it is worth mentioning that given the high-payoff for Row if the combination *Top-Left* was played and the fact that 58.0% of Column players in our experiment chose *Left*, the higher expected payoff for Row is obtained if *Top* is chosen (more than three times the expected payoff for *Bottom*). On the other hand, for the Column Player the optimal choice, given the actual play frequencies of Row in our study, was *Right*.

Table 3.2

Choices and expected payoffs

Player	Choice	Expected Payoff *
Row	*Top*	202.5
	Bottom	56.8
Column	*Left*	44.4
	Right	75.6

*Experimental currency units

On the other hand, if we divide the sample by playing order, that is to check the effect of having played the other's player role before, we find a marginally significant difference between distributions for Column choices ($p<.10$). That is, when playing as Column, those who played first as Row, choose more often *Right* which un-matches Row's high own-payoff choice *Top.* Therefore it seems that even with no feedback, strategic reasoning might be positively affected at an introspective level.

Table 3.3

Choices and playing order effects

Player		1st time	2nd time	Chi2	p-value
Row	[Top; Bottom]	[91.1%; 8.9%]	[86.1%; 13.9%]	0.51	0.48
Column	[Left; Right]	[69.4%; 30.6%]	[48.9%; 51.1%]	3.47	0.06

3.4.2 Cognitive abilities and choices

Below we report the effect of the measured cognitive variables on the probability of choosing one of the pure strategies (*Top* for Row and *Right* for Column). We do this by implementing a Probit specification that measures the effect of each cognitive attribute on the latent propensity for a positive result. In Table 3.4 we report two reduced Probit models, the first one for Row choices and the second for Column choices. Both reduced models are obtained using the backward stepwise selection method, that is, starting in each case with the full model (including all the variables) and removing one by one each variable where p>.10. Full models are reported in Appendix (5.2.1).

Table 3.4

Probit regressions. Cognitive abilities and choices

Cognitive ability	Variable	Dependent variable:	
		Row chose *Top*	Column chose *Bottom*
Non-verbal reasoning	RAVEN_P	**-0.018 *** (0.010)	
Working memory	3-BACK_SCORE		**-0.120 **** (0.059)
Flexibility of thinking	WIS_PPERSM	**0.020 **** (0.009)	
Response inhibition	SS_RT	**0.014 **** (0.006)	
	DUMMY_ROW		**-0.931 *** (0.484)
	CONSTANT	**-1.758** (1.406)	**-1.758** (1.406)
Observations		81	81
Log likelihood		-21.797	-51.067
χ^2 (n)		12.220	8.060
Prob > χ^2		0.005	0.018
McFadden R-squared		0.229	0.073

*10% confidence level, **5% confidence level, ***1% confidence level.
Standard errors in parentheses.
For X^2, n=3 for Row choice model and n=2 for Column choice model.

From the first Probit regression in Table 3.4 we can infer that among the four cognitive aspects of player that we assessed, three of them (reasoning ability, flexibility of thinking and response inhibition) are relevant to explain the propensity of Row Players to choose *Top*. In particular the probability of choosing the high own-payoff option, *Top* increases with: a decrease in reasoning ability (Raven progressive matrices percentile); with an increase in cognitive flexibility (Wisconsin CST percentile) and a decrease in response inhibition (Stop-Signal Reaction Time Task). Working memory ability doesn't seem to play any role for Row choices.

On the other hand, for Column choices, subjects with greater working memory (3-Back task score) and those who previously played as Row are less likely to choose option *Right*, which un-matches the high other-payoff *Top*[27].

3.4.3 Gender effects

Although our experiment was not designed to study gender effects, we find some evidence of differentiated behavior across genders. In the light of these findings we considered necessary to dedicate a part of the essay to reassess our results. The conclusions we obtain are constrained by the limitations imposed by our data, which, as a consequence of not having been conceived to study this aspect, has a strongly unbalanced design (28 males and 53 females).

Despite these limitations, we find that for Row choices, the difference between the observed distributions of pure strategies depending on the gender is statistically significant ($p<0.05$) while for Column choices the difference across genders is not statistically significant.

Table 3.5
Choices and gender effects

Player		Males	Females	Chi2	p-value
Row	[Top; Bottom]	[100%; 0%]	[83.0%; 17.0%]	5.35	0.02
Column	[Left; Right]	[64.3%; 35.7%]	[54.7%; 45.3%]	0.69	0.41

If we focus on Row choices and divide the sample by playing order, we find that although choice distributions for each gender are almost identical across orders, the gender difference is only significant for those who played Row first ($p<0.10$). This result is due to the fact that among those who played Row first the number of males and females was similar (20 and 25, respectively) while very few male participants (8 vs. 28 females) played Column first, which undermines the power of statistical inference.

[27] Marginal effects of the variables are small, but significant in all cases $p<0.10$.

Table 3.6

Row player choices by playing order and gender

Order	Row choice	Males	Females	Chi2	p-value
Row 1st	[Top; Bottom]	[100%; 0%]	[84.0%; 16.0%]	3.51	0.06
Row 2nd	[Top; Bottom]	[100%; 0%]	[82.1%; 17.9%]	1.66	0.20
Chi2		-	0.02		
p-value		-	0.90		

On the other hand, within genders we find absolutely no evidence of order effects (as we did not find in the aggregated data). Consequently there is no interaction between gender and order in our sample. Therefore, although we cannot generalize our results to include any play order, there seems to be evidence that male player’s have a greater propensity to select the high own-payoff choice. If this conclusion is correct, this could be due either to gender differences in unmeasured idiosyncratic factors, such as risk aversion, or due to differences in the cognitive ability levels across genders. Using diverse statistical approaches we reject the later[28]. This has two consequences: on one hand, possible gender differences in choices could be, to some degree, driven by unmeasured individual characteristics and; on the other, results regarding the effect of cognitive abilities on choices are not driven by gender differences.

Row Players: To verify this affirmation and to see if the results we obtained in the previous section were not driven by gender differences we also re-estimate the Row Probit model only for females[29]. In Table 3.7 we report the reduced Probit model including only relevant independent variables (full models in Appendix 5.2.1). Our previous results hold when only females are considered, that is, relevant variables and signs are the same, while magnitudes of the coefficients are very close to the estimations with the full sample.

[28] We applied the Mann Whitney U-test for each cognitive ability variables across gender; we also estimated Probit regressions including all cognitive variables and gender as dependent variable. We didn’t find any significant result (p>0.10).

[29] Notice that we cannot include a dummy for gender, since gender creates a perfect separation of Row player choices that is because 100% of male players choose *Top*.

Table 3.7

Probit regression: Females' cognitive abilities and choices

Cognitive ability	Variable	Dep. Var: Row chose *Top*
Non-verbal reasoning	RAVEN_P	**-0.020** * (0.011)
Flexibility of thinking	WIS_PPERSM	**0.022** ** (0.010)
Response inhibition	SS_RT	**0.015** ** (0.006)
	CONSTANT	**-2.226** (1.520)
Observations		53
Log likelihood		-21.797
χ^2 (n)		12.410
Prob > χ^2		0.006
McFadden R-squared		0.257

*10% confidence level, **5% confidence level, ***1% confidence level.
Standard errors in parentheses.

Column players: With respect to Column choices (Table 3.8) there is no significant difference between the frequencies of choices across genders, even taking into account possible order effects. On the other hand, we find that female Column Players that previously played as Row, choose more often *Right* (60.0%) than those playing first as Column (32.1%) and the difference is significant ($p < 0.05$). For males, we observe the same shift, but less intense (40% vs. 25%) and statistically non significant, due to both less intensity and the fact that the sample is smaller and highly unbalanced, which undermines the power of statistical inference.

Table 3.8

Column player choices by playing order and gender

Order	Column choice	Males	Females	Chi2	p-value
Column 1st	[Left; Right]	[75.0%; 25.0%]	[67.9%; 32.1%]	0.15	0.70
Column 2nd	[Left; Right]	[60.0%; 40.0%]	[40.0%; 60.0%]	1.78	0.18
Chi2		0.56	4.14		
p-value		0.45	0.04		

Going back to our Probit specification, if we include the variable gender, the corresponding coefficient is not significantly different from zero ($p > 0.10$). The same applies if we include an interaction term between gender and order. Neither gender, nor order or the interaction coefficients are significantly different from zero ($p > 0.10$). Therefore, order is only significant for the aggregated sample. On the other hand, when we estimate separate regressions for each gender we find that the best model for females includes only the order dummy variable, while for males no explanatory variable is significant. In both cases, the working memory measure (3-Back) is excluded due to the low significance Nevertheless, the sign and the magnitude of the coefficient for both genders is similar.

3.5 Concluding remarks

In the framework of an asymmetric matching pennies game we have shown that subjects with lower inhibition and reasoning capacity are more likely to be attracted by a salient payoff, thus becoming more predictable and easier to exploit by other players. On the opposite side of the game, players with a higher working memory are more likely to exploit the aforementioned pattern of behavior. Generally speaking, our findings are novel in that we shed some light on the under-investigated issue of why the predictions of a mixed strategy equilibrium may fail to be verified by empirical distributions generated as a population of one-shot choices in a game lacking pure strategy equilibria.

3.6 References

Anderson, S. P., Goeree, J. K., & Holt, C. A. (2002). The logit equilibrium: A perspective on intuitive behavioral anomalies. *Southern Economic Journal*, *69*(1), 21–47.

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, *55*(1), 1–18.

Azar, O. H., & Bar-Eli, M. (2011). Do soccer players play the mixed-strategy Nash equilibrium? *Applied Economics*, *43*(25), 3591–3601.

Baddeley, A. (1992). Working Memory. *Science*, *255*(5044), 55.

Boylan, R. T., & Grant, S. (2008). Incorporating fairness in generalized games of matching pennies. *International Journal of Economic Theory*, *4*(4), 445–458.

Brown, J. N., & Rosenthal, R. W. (1990). Testing the minimax hypothesis: a re-examination of O’Neill's game experiment. *Econometrica*, *58*(5), 1065–1081.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2002). *A cognitive hierarchy theory of one-shot games: Some preliminary results* (Working paper, California Institute of Technology). Retrieved from: http://www.hss.caltech.edu/~camerer/thinking2002.pdf

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A Cognitive Hierarchy Model of Games. *Quarterly Journal of Economics*, *119*(3), 861-898.

Cañas, J., Quesada, J., Antolí, A., & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. *Ergonomics*, *46*(5), 482-501.

Chiappori, P. A., Levitt, S. D., & Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, *92*(4), 1138-1151.

Costa-Gomes, M., & Crawford, V. P. (2006). Cognition and Behavior in Two-Person Guessing Games: An Experimental Study. *American Economic Review*, *96*(5), 1737-1768.

Crawford, V. P., & Iriberri, N. (2007). Fatal attraction: Salience, naivete, and sophistication in experimental “Hide-and-Seek” games. *American Economic Review*, *97*(5), 1731–1750.

Eichberger, J., & Kelsey, D. (2011). Are the treasures of game theory ambiguous? *Economic Theory*, *48*(2-3), 313-339.

Gallice, A. (2007). *Best Responding to What? A Behavioral Approach to One Shot Play in 2x2 Games* (Munich Discussion Paper 2007-6). Germany: Department of Economics University of Munich. Retrieved from: http://epub.ub.uni-muenchen.de/1365/

Goeree, J. K., & Holt, C. A. (2001). Ten Little Treasures of Game Theory and Ten Intuitive Contradictions. *American Economic Review*, *91*(5), 1402-1422.

Goeree, J. K., & Holt, C. A. (2004). A model of noisy introspection. *Games and Economic Behavior*, *46*(2), 365-382.

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2003). Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, *45*(1), 97–113.

Gonzalez, C., Thomas, R., & Vanyukov, P. (2005). The relationships between cognitive ability and dynamic decision making. *Intelligence*, *33*(2), 169-186.

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404-411.

Harsanyi, J. C. (1973). Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, *2*(1), 1–23.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sort Test Manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources, Inc.

Hsu, S. H., Huang, C. Y., & Tang, C. T. (2007). Minimax Play at Wimbledon: Comment. *American Economic Review*, *97*(1), 517–523.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, *55*(4), 352-358.

Kovash, K., & Levitt, S. D. (2009). *Professionals do not play minimax: Evidence from major league baseball and the national football league* (NBER Working Paper 15347). Cambridge, MA: National Bureau of Economic Research. Retrieved from: http://www.nber.org/papers/w15347.pdf

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, *10*, 6-38.

McKelvey, R. D., Palfrey, T. R., & Weber, R. A. (2000). The effects of payoff magnitude and heterogeneity on behavior in 2×2 games with unique mixed strategy equilibria. *Journal of Economic Behavior & Organization*, *42*(4), 523-548.

Nash, J. (1951). Non-Cooperative Games. *The Annals of Mathematics*, *54*(2), 286-295.

Ochs, J. (1995). Games with unique, mixed strategy equilibria: An experimental study. *Games and Economic Behavior*, *10*, 202-217.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*, 46-59.

O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the National Academy of Sciences*, *84*(7), 2106-9.

Palacios-Huerta, I. (2003). Professionals Play Minimax. *Review of Economic Studies*, *70*(2), 395-415.

Pickering, S. J., Baqués, J., & Gathercole, S. E. (1999). *Bateria de tests de memòria de Treball* (mimeo). Barcelona: Laboratori de Memòria del U.A.B.

Rapoport, A., & Boebel, R. B. (1992). Mixed strategies in strictly competitive games: A further test of the minimax hypothesis. *Games and Economic Behavior*, *4*(2), 261–283.

Rosenthal, R. W., Shachat, J. M., & Walker, M. (2003). Hide and seek in Arizona. *International Journal of Game Theory*, *32*(2), 273-293.

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime: A User's Guide*. Pittsburgh: Psychology Software Tools, Inc.

Shachat, J. M. (2002). Mixed Strategy Play and the Minimax Hypothesis. *Journal of Economic Theory*, *104*(1), 189-226.

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in cognitive sciences*, *12*(11), 418-24.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Walker, M., & Wooders, J. (2001). Minimax play at Wimbledon. *American Economic Review*, *91*(5), 1521-1538.

Wooders, J. (2010). Does Experience Teach? Professionals and Minimax Play in the Lab. *Econometrica*, *78*(3), 1143-1154.

Essay 3

4. The Lottery-Panel Task for Bi-dimensional, Parameter-Free Elicitation of Risk Attitudes: Implementation and Results[30]

4.1 Introduction

Human beings are usually acting in different contexts and environments. Each individual expresses needs, preferences, attitudes, and ideologies through different actions in each of the domains in which he or she chooses or simply happens to be. As contexts become closer or somehow related to each other, actions by the same individual should also become more related in one way or another. In fact, in an ideal world in which a subjects' personality is a compact and stable system of values and idiosyncratic features, behavior in related contexts should confirm the revelation of the same person. Based on this idea, social scientists like psychologists and economists often try to explain behavior heterogeneity through idiosyncratic differences across subjects. Such differences are usually captured by exposing individuals to decision making tasks or attitude questionnaires. Famous examples are a plethora of intelligence tests and personality inventories used by the psychologists to asses an individual's performing skill and propensity to one or another type of action.

In order to produce reliable tests, it is necessary to invest a substantial amount of effort in (i) developing the task and proposing it to the scientific community, (ii) standardizing the format and applying it among large populations, (iii) generating result distributions by subject category and (iv) identifying successful tasks as reliable approximations of an idiosyncratic factor. Moreover, the search of associations among decisions in different tasks is a main motivator for experimentalists. For example, when studying the effects of intelligence on complex decision making, psychologists correlate scores in, say, Raven's (1976) Advanced Progressive Matrices (APM), and performance in complex microworlds, like NEWFIRE or COLDSTORE (Rigas et al., 2002). Beyond the question of what explains what, a systematic rejection of such associations would

[30] Part of this essay is based on a joint paper with Aurora García-Gallego, Nikolaos Georgantzís and Ainhoa Jaramillo-Gutiérrez entitled "The lottery-panel task for bi-dimensional parameter-free elicitation of risk attitudes", published in Revista Internacional de Sociología (RIS), Special Issue on Behavioral and Experimental Economics. Vol. 70 (1): 53-72, March 2012. We acknowledge financial support by *Bancaixa* (P1-1B2010-17) and the Spanish Ministry of Science and Innovation (SEJ2008/04636/ECON).

confine experimental results to the specific setting in which they were obtained, undermining the practical relevance of the research outside the lab. This process is parallel and significantly synergic to the very important endeavor of producing correct theories on the measured aspect itself. However, metaphorically speaking, looking for appropriate tasks in the absence of a perfect theory is like the practice in medicine of establishing clinical protocols for the cure of a disease even before the disease is fully understood.

Paradoxically, economists have failed so far to agree upon the systematic use of a stable task eliciting individual attitudes towards monetary uncertainty. Even the need for external risk measurements is often not recognized by some economists, who frequently explain the effect of risk preferences on observed behavior by theoretically deriving the sufficient conditions for this effect to emerge, thus explaining fact Y by its sufficient (but not necessary) condition X (e.g., Campo, Guerre, Perrigne, & Vuong, 2011; Cox & Oaxaca, 1996; Goeree, Holt, & Palfrey, 2002). Furthermore, in the few cases in which a task has been used more often than others, this has not been done in a systematic way, creating small non-comparable samples of observations. Thus, experience and statistical significance have not been built in a cumulative way. Even worse, as we argue below, the so called risk attitude tests are ignoring past evidence and new theories on individual behavior in risky contexts.

In this essay, a decision making task which is specific to individual decision making in contexts involving uncertainty of the monetary consequences is proposed. Given the importance of uncertainty in modern societies exposed to macroeconomic financial shocks, linking individual attitudes towards risk with actions in other domains would give us a powerful tool to assess the role of personality traits on market functioning. The task discussed in the following pages provides a context for the elicitation of risk attitudes in a way that is both compatible with the need for a multidimensional assessment and robust to alternative mathematical specifications and parameterizations of the model used to organize the data.

The remaining of the essay is structured as follows: Next Section reviews economic theories of risky decision making and comments on some devices used to elicit risk attitudes as an external explanatory factor of behavior in other contexts. Study 1 reports results obtained from the application of the SGG lottery-panel test by Sabater-Grande and Georgantzis (2002). Study 2 proposes and applies experimentally three new versions of the task designed to assess both the effect of large stakes and mixed outcome lotteries on risk attitudes. The last section summarizes results and presents general conclusions on both studies.

4.2 Economic and psychological theories and tests of risk attitudes

It is well known that individuals faced with a probability *p* to earn a given amount of money *x* may be willing to pay less or more than the product *p·x* to earn access to this possibility. From a mathematical point of view the product *p·x* should be used as a certainty equivalent of the aforementioned lottery. Thus, if the probability *p* of earning *x* Euros were evaluated by its mathematical expectation all people would accept to pay less and would reject to pay more than a certain amount of *p·x* Euros in order to participate in the lottery. But, as we know people are not mathematical machines nor identical problem-solving automata.

An early explanation of why subjects do not evaluate risky choices by their mathematical expectations is attributed to the Expected Utility Theory (EUT) (von Neumann & Morgenstern, 1944). According to the theory, when comparing a lottery $L_1 = (p_{11}, x_{11}€;...p_{1n}, x_{1n}€)$ with $L_2 = (p_{21}, x_{21}€;...p_{2m}, x_{2m}€)$, where p_{ji} is the probability that the i th best outcome of lottery j occurs, yielding a reward of x_{ij} €, an agent whose utility is $U(x)$, with $U'(*) > 0$, will strongly prefer L_1 to L_2, as long as

(1) $$\sum_{i=1}^{n} p_{1i} \cdot U(x_{1i}) > \sum_{i=1}^{m} p_{2i} \cdot U(x_{2i}) .$$

The preference for less risky projects is then explained by a negative second derivative of $U(x)$, implying a decreasing marginal utility from money, a condition often used as synonymous to *risk aversion*. Despite its survival as the main paradigm in economics as observed by Rabin and Thaler (2001), the EUT was proved to be an incorrect descriptive model since Allais' (1953) paradox, emerging when subjects are faced to alternative lottery pairs with same probability/reward ratios. According to (1), such lotteries should be ranked in the same way, whereas people systematically change their choice in favor of the certain payoff when this becomes part of the feasible set. Kahneman and Tversky (1979) proposed an alternative model, Prospect Theory (PT), assuming that people implicitly use non linear weights $w(p)$ to evaluate probabilities. Therefore, in our example, L_1 would be strongly preferred to L_2, if:

(2) $$\sum_{i=1}^{n} w(p_{1i}) \cdot U(x_{1i}) > \sum_{i=1}^{m} w(p_{2i}) \cdot U(x_{2i})$$

That is, not only the outcomes create non linear utility responses but also probabilities are distorted in the decision maker's mind. Therefore, new possibilities emerge concerning what we could expect from a rational decision maker's actions. Consequently, PT accommodates Allais' paradox, whereas it reduces to EUT for $w(p) = p$. Also, observing that losses and gains are processed differently, Tversky and Kahneman (1992) assumed later a power utility function defined separately over gains and losses: $U(x) = x^a$ if $x > 0$, and $U(x) = -\lambda(-x)^b$ for $x < 0$. So a and b are risk aversion parameters, and λ is the coefficient of loss aversion. This new version, called Cumulative Prospect Theory (CPT), defines probability weighting over the cumulative probability distributions, offering an explanation of risk-loving behavior for payoffs below their reference point (losses), while exhibiting risk-averse behavior for rewards above their reference point (gains). The form of the probability weighting function proposed by Tversky and Kahneman (1992) has been widely used for both separable and cumulative versions of PT, and assumes weights $w(p) = p^\gamma / \left[p^\gamma + (1-p)^\gamma\right]^{1/\gamma}$. Therefore, in its simplest formulation, CPT explains risk attitudes using a minimum of four parameters, a, b, λ and γ.

Despite the fact that PT and its improved version CPT has become the leading model to organize behavior under risk, there is a growing body of theories that introduce new elements, such as the one proposed by Birnbaum and Navarrete (1998) which can explain violations of stochastic dominance by introducing a third component of risky decision making, namely the attention paid by subjects to the best outcomes among those feasible in a given lottery.

Our overview does not pretend to narrate the history of economic theories of decision making, in fact we have intentionally omitted heuristics and other theories which cannot be used to propose tasks for the elicitation of risk attitudes, we simply want to stress the fact that the evolution of these theories achieves the aim of accommodating phenomena which invalidated earlier theories by the use of more degrees of freedom.

Contrary to this evolution of theories towards more complete and complex descriptions of human behavior in risky environments, all tests currently used are fundamentally uni-dimensional, despite their creation in the post-PT era. This does not mean that all studies of behavior under uncertainty have ignored the multi-dimensional approach dictated by modern theories. In fact, a fruitful line of research has specifically designed and analyzed experimental data to estimate parameters for utility and probability weighting functions, such as the Tversky and Kahneman (1992) probability weighting function and other specifications like, for example, Goldstein and Einhorn (1987) and Prelec's (1998) two-parameter specification. Furthermore, the nonlinearity of responses to probabilities has even been confirmed at the level of neural responses (M. Hsu, Krajbich, Zhao, & Camerer, 2009), and, for aversive outcomes (Berns, Capra, Chappelow, Moore, & Noussair, 2008). However, in order to produce ready-to-use data, the elicitation of risk attitudes as an explanatory factor of behavior in another context should not depend on the parameterization or even the theory used. Mapping choices on parameters of utility and probability weighting functions is further complicated by the observation that we may even have to switch between theories in order to account for the heterogeneity observed (Harrison, Rutström, & Towe, 2009).

In many recent economic studies, a measure of risk aversion is obtained by the use of the Holt and Laury (2002) HL procedure. Although the task was not, initially, proposed as an external risk-related task to explain behavior in other contexts, it has served this purpose in several occasions (e.g., Andersen, Harrison, Lau, & Rutström, 2008; Goeree, Holt, & Palfrey, 2003; Harrison & List, 2007; Lusk & Coble, 2005). Due to its uni-dimensionality, costlessly allowing a one-to-one mapping of choices on specific utility parameters, the test entails a possible loss of information due to under-specification of risk attitudes, which is also likely to reduce its power to explain behavior in other contexts. This is also true for the whole set of alternative procedures used by economists to elicit risk attitudes (e.g., Abdellaoui, 2000; Abdellaoui, Bleichrodt, & Paraschiv, 2007; Bleichrodt, Pinto, & Wakker, 2001; Camerer & Ho, 1994; Carbone & Hey, 2000; Hey & Orme, 1994; Stott, 2006; P. Wakker & Deneffe, 1996).

The HL task elicits one individual *datum* from each block of 10 binary choices, designed to obtain the switching point from a less risky to a more risky alternative. This causes a practical problem since some choices do not satisfy the “single-switching” condition. Posterior applications have opted for different solutions to this problem, leading to a variety of alternative implementations which, together with the plethora of designs aimed at identifying other biases of the set up, have created an –undesirable, for our purposes– plethora of non comparable datasets. Contrary to the problem of non comparability among small data sets, several studies (e.g., Wang, Rieger, & Hens, 2011; E. U. Weber & Hsee, 1998, 1999) use hypothetical simple questions among large and even international samples, which however have not been used to explain behavior in other contexts.

A broadly used test among psychologists is Zuckerman's (1978) Sensation Seeking Scale (SSS) with which our task exhibits some correlation (Georgantzís, Genius, García-Gallego, & Sabater-Grande, 2003). SSS is structured as a YES-NO questionnaire on attitudes towards risky activities under four subscales separating subject’s riskiness in different domains, none of which is strictly speaking financial. The economic domain of risk is used in the Iowa Gambling Task (Bechara, Damasio, Damasio, & Anderson, 1994). The task was originally aimed at measuring a subject’s difficulty to identify the most profitable deck, from which he or she should, thereafter, extract all cards. Using the task as an external risk attitude elicitation device implies significant loss of control, because it mixes risk preferences with a subject’s learning ability (a “slow” learner can be confused with a risk loving subject or one with low levels of loss aversion) and it does not fully account for different learning histories. For space reasons, we will not review other tests occasionally used to elicit risk attitudes as an explanatory factor of behavior in other contexts. Rather, we will risk a generalization. All existing tasks suffer from either lack of systematic replication in a stable format generating statistics with large comparable datasets, or they are insufficiently justified as measures of risk attitudes isolated from other parallel phenomena. Furthermore, they are all uni-dimensional.

4.3 Study 1: The SGG lottery-panel task

4.3.1 Task overview

The SGG lottery-panel task was originally used to study risk preferences parallel to cooperation/competition in prisoner's dilemma games. Riskier subjects were found to be more cooperative. The task consists of four different panels, like those in Figure 4.1, every one of which contains ten different lotteries. In each lottery, subjects can win a payoff (x) with a probability (p) and otherwise nothing.

Panel 1

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	1.00	1.10	1.30	1.50	1.70	2.10	2.70	3.60	5.40	10.90
Choice			x							

Panel 2

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	1.00	1.20	1.50	1.90	2.30	3.00	4.00	5.70	9.00	19.00
Choice				x						

Panel 3

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	1.00	1.70	2.50	3.60	5.00	7.00	10.00	15.00	25.00	55.00
Choice							x			

Panel 4

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	1.00	2.20	3.80	5.70	8.30	12.00	17.50	26.70	45.00	100.00
Choice									x	

Figure 4.1 *The SGG lottery-panel task and example of subject choices.*

Subjects choose (marking the preferred lottery as in the example of Figure 4.1) one of the ten lotteries from each panel. In the implementation of the task with real money, one of these four panels, selected randomly at the end of the session, is used to determine a subject's earnings in the experiment. The range of winning probabilities in all panels is the same (from 1 to 0.1 in steps of 0.1). The payoff associated to each lottery's winning probability is constructed using the rule:

(3) $$E(L_{ij}) = p_{ij} \cdot x_{ij} = c_j + (1 - p_{ij}) \cdot t_j \Rightarrow x_{ij} = {}^{c_j + (1-p_{ij}) \cdot t_j}\!\big/_{p_{ij}}.$$

$E(L_{ij})$ is the expected value of lottery L_{ij}, where $i \in \{1,2,...,10\}$ designates one of the 10 lotteries offered in panel $j \in \{1,2,3,4\}$. The parameter c_j is a constant amount of money which is fixed for this dataset to 1€. The parameter $t_j \in \{0.1,1,5,10\}$ is a panel-specific risk premium, which generates an increase in the lotteries' expected values as we move from safer to riskier options within the same panel. All the panels begin with a sure amount of 1€, which is increased as winning probabilities are decreased, resulting in increments of expected values as we move from left to right within each panel. These increments are larger as we move from panel 1 to panel 4. This structure implies that more risk-averse subjects choose lotteries closer to the left of a panel.

Intuitively, this test exposes subjects to the entire range of probabilities and a systematically generated spectrum of monetary rewards from 1 Euro to the relatively high payoff of 100 Euros. At the same time, the test offers a range of different returns to risk so that a more risk averse subject might refuse to take risky options in the first or the second panel, but could be attracted to risky prospects when a high return is offered in panel 3 and 4. Thus, unlike all unidimentional tests, this task may be used to classify subjects not only according to their willingness to take risks, but also with respect to their propensity to change across different risk-return combinations. This idea is further developed in the following pages.

In terms of EUT, a subject with constant relative risk aversion (CRRA), as implied in the utility function $U(x) = \frac{x^{1-r}}{1-r}$ makes choices which associate higher risk aversion parameters r to safer choices in each panel, moreover, for a given risk aversion parameter, weakly monotonic transitions towards riskier choices are predicted as we move from panel 1 to panel 4 (García-Gallego, Georgantzís, Navarro-Martínez, & Sabater-Grande, 2011). All risk neutral and risk loving subjects should choose the lotteries at the far right extreme of the panels.

Considering the fact that with 4 choices the researcher obtains 4 different observations (as opposed to 10 choices for 1 observation in HL) per individual subject, we can easily see that the test parsimoniously produces a panel rather than a single column of data. By definition, this corresponds to a multi-dimensional description of individual attitudes towards risk.

4.3.2 Implementation and results

Since its first implementation, the SGG test has been used in several occasions producing various small experimental datasets (e.g., Brañas-Garza, Georgantzís, & Guillén, 2007; Brañas-Garza, Guillén, & López del Paso, 2008; García-Gallego et al., 2011; Georgantzís et al., 2003). Here, we report results from a large dataset (N=785), obtained between 2003 and 2008, at the Laboratorio de Economía Experimental (Universitat Jaume I, Castellón-Spain) under comparable conditions, paying special attention to the bi-dimensional nature of decision making and its implications for the explanation of behavior in other contexts. Figure 4.2 depicts the frequency of choices when all data from all panels are pooled together. Given the variation in prizes and payment methods, this image corresponds to what could be seen as a randomized experiment over the probability space. The peak on the certain payoff captures a certainty effect. A peak on the other extreme (*p=0.1*) as well as a valley on *p=0.9* are both compatible with over-(under-) weighting of small (large) probabilities predicted in PT. Strong attraction of choices towards the "center" (*p=0.5*) may be the result of subjects' familiarity with the *p=½* probability or simply because of an embedding bias similar to that reported by Bosch-Domènech and Silvestre (2006) on HL.

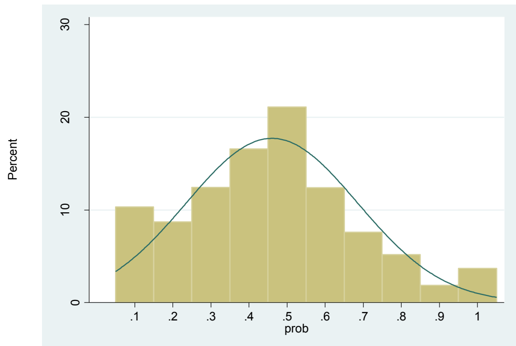

Figure 4.2 *Histogram of subjects' pooled probability choices across all panels and implementation conditions.*

In Figure 4.3 we present the same dataset broken down by panel, gender and reward method (hypothetical, N=384; real, N=401). Males are less risk-averse than females. However, males and females behave in more different ways when playing hypothetical lotteries than real ones. Actually, with real rewards, mean choice varies significantly across genders only in panel 3 and 4 (2.7 and 3.9 percentage points at 5% and 1% confidence level, respectively). Responsiveness to risk-premium increases, captured by choice variation across panels, is similar for males and females. Specifically, when faced with hypothetical payoffs, both males and females make less risk-averse choices, the higher the reward, while, counterintuitively, when playing with real payoffs, riskier choices are observed in panels with lower risk-returns.

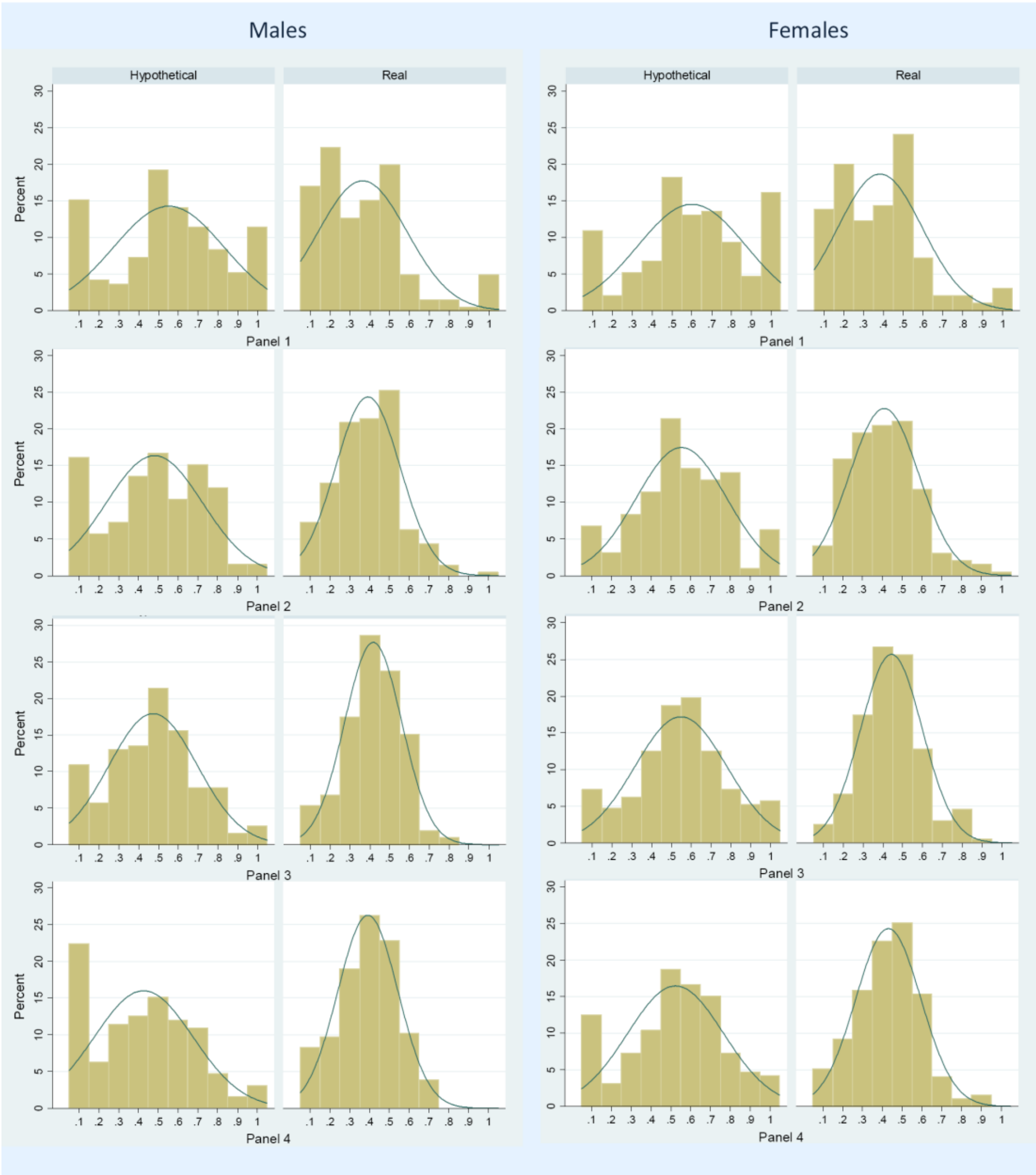

Figure 4.3 *Histograms of subjects' probability choices by panel, implementation conditions and gender.*

We have argued that it should be a main concern for experimentalists and decision theorists whether a subject's decision under one condition meaningfully relates to behavior under another condition.

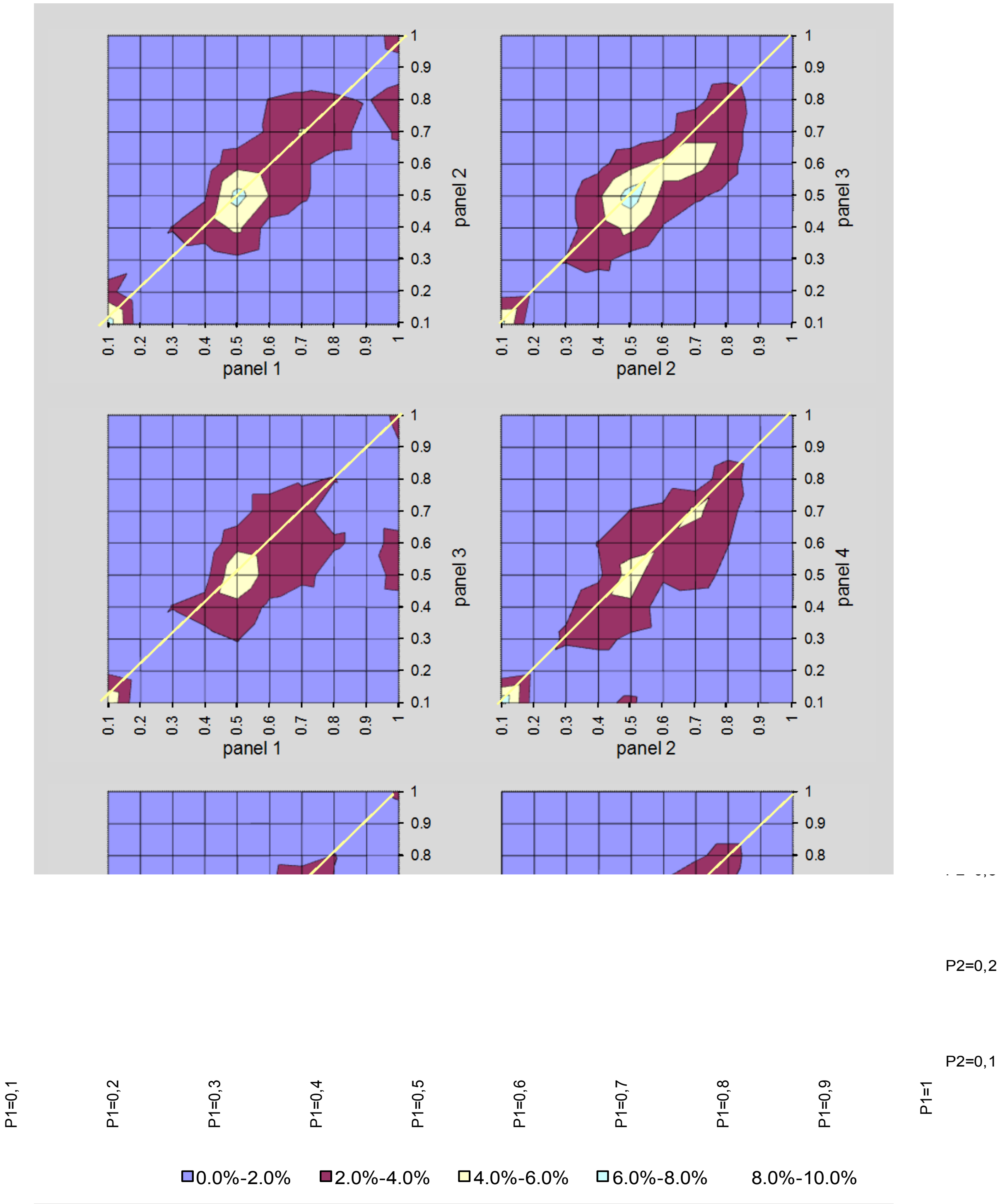

Figure 4.4 *Subject's choices across panel pairs for hypothetical payoff lotteries. Legend percentage ranges refer to proportion of subjects choosing combinations indicated in each chart label.*

Figures 4.4 and 4.5 present an aspect of behavior which is missed by other tests. Each graph presents the joint density of individual choices across panel pairs. Each color represents a percentage, i.e. the proportion of subjects whose choice combinations in each panel pair correspond to that specific chart label. Lower risk taking in one panel predicts a lower risk taking in any other panel and, at the same time, reactions to the variation of risk returns across different panels seem to be rather moderate.

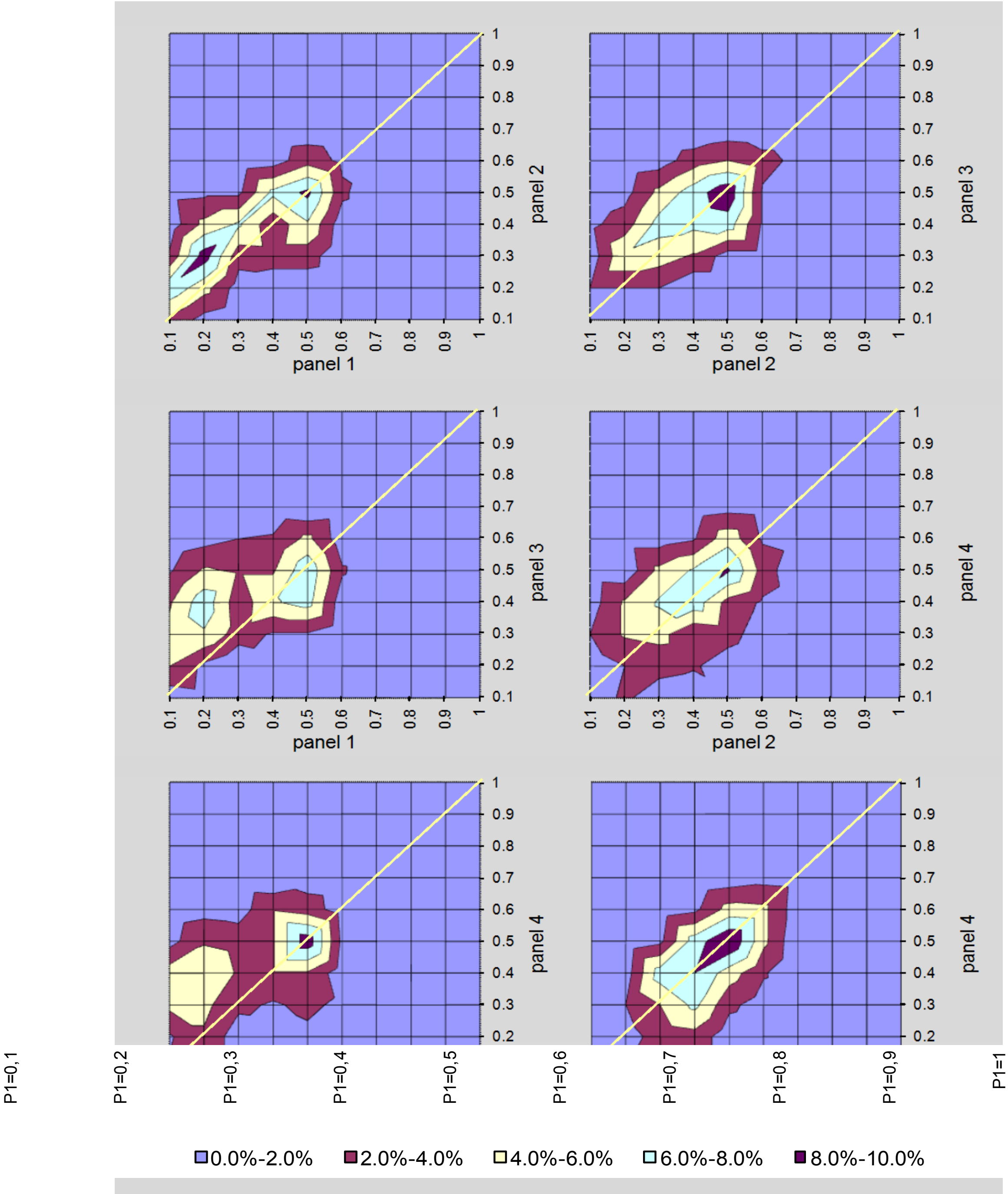

Figure 4.5 *Subjects' choices across panel pairs for real payoff lotteries. Legend percentage ranges refer to proportion of subjects choosing combinations indicated in each chart label.*

As expected, reactions are more visible across more "distant panels", showing that a bigger shock is necessary to guarantee a change of choices. This within-subjects pattern reproduces in a more reliable way what we have already observed, namely, that the use of real rewards makes subjects to switch to safer options in the presence of higher returns to risk.

4.3.3 Principal Component Analysis

It is clear that multidimensional descriptions of risk attitudes require obtaining more than one choice per individual. This is done by the SGG test through the use of the four panels. However we have not shown yet that, first, the additional information obtained significantly improves the description of behavior and, second, that this improvement leads to a higher power of our task to explain behavior in other contexts.

We use Principal Component Analysis (PCA) to construct two synthetic variables (the first two components) capturing 85% of subjects' choice variance. These variables have the following advantages: (1) they are subject to economic interpretation and, (2) since they are by construction orthogonal among each other, they can be used as explanatory variables of the same econometric model. Intuitively, the first component can be interpreted as an arithmetic mean of choices across the four panels given that the loads of each panel in this component are similar and of the same sign. The second component involves a juxtaposition of panels 1 and 2 on one hand and 3 and 4 on the other, which can intuitively be seen as a measure of sensitivity to risk-premium variations.

As observed in Table 4.1, the component is loaded more by the extreme panels 1 (negatively) and 4 (positively) than by choice differences across the adjacent panels, 2 and 3. Intuitively, the first component is increasing in the average probability of the lottery chosen in the four panels and can be seen as a standard measure of risk aversion. The second component can be seen as a measure of a subject's sensitivity to variations in the return to risk in the direction of higher risk taking in the presence of higher returns to risk. While this confirms our comments on Figures 4.4 and 4.5, it provides a formal motivation for the use of bi-dimensional descriptions of risk attitudes, summarized as individual choice averages and choice variability across contexts (panels).

Table 4.1

Principal Components Analysis

Component	Eigenvalue	Percentage (%)	Cumulative %
Comp. 1	2.742 ***	68.54	68.54
Comp. 2	0.670 ***	16.75	85.29
Comp. 3	0.307 ***	7.67	92.96
Comp. 4	0.282 ***	7.04	100
	Panel	**Coefficient**	**Std. Error**
Comp. 1			
	Panel 1	0.489 ***	0.016
	Panel 2	0.517 ***	0.013
	Panel 3	0.521 ***	0.013
	Panel 4	0.472 ***	0.017
Comp. 2			
	Panel 1	0.577 ***	0.029
	Panel 2	0.372 ***	0.035
	Panel 3	-0.317 ***	0.036
	Panel 4	-0.654 ***	0.027

*** significant at 1% level of confidence.

Using these two components we reconsider gender and hypothetical/real reward effects. It can be seen on Figure 4.6 that gender differences are specific to the first component, while they diminish or even vanish in the second component. Therefore, males are less risk averse than females but both genders are similar in terms of their sensitivity to variations in the return to risk. Regarding differences between hypothetical and real rewards, both components are relevant. According to the first component, subjects make safer choices in hypothetical lotteries, while, according to the second component they switch more across panels with real rewards, but opposite to the expected pattern of riskier choices for higher risk-returns.

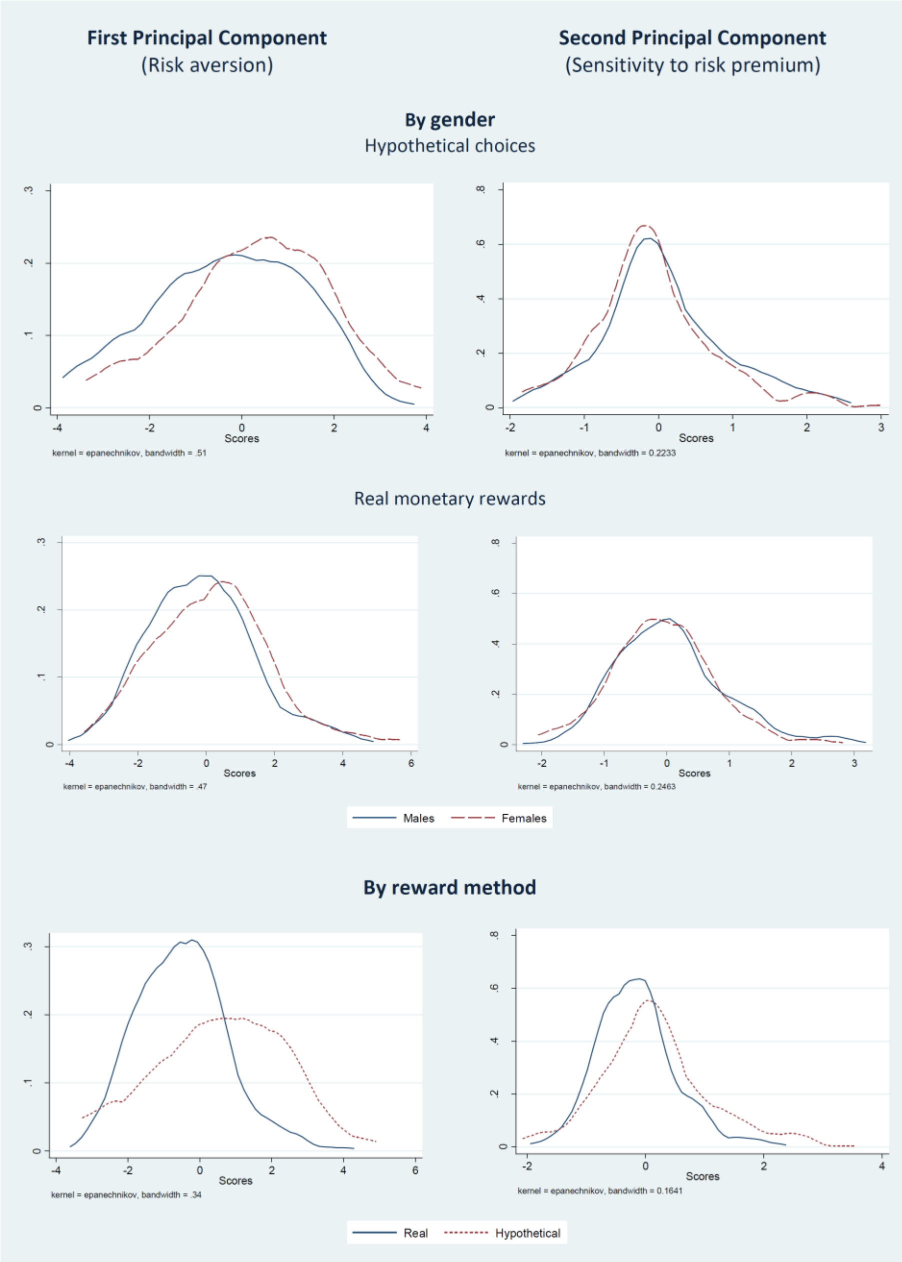

Figure 4.6 *Kernel density estimates for first and second component scores, by gender and reward method.*

3.3.4 Using the SGG test to explain behavior: An example.

García-Gallego, Georgantzís, Pereira, and Pernías-Cerillo (2005) conducted experiments on pricing where firms have some captive clients and they also compete for informed consumers using price comparisons on the Internet. During 50 periods, subjects face the dilemma of setting high prices to benefit from captive clients or lower prices to compete for informed consumers too. Parallel to the main experiment controlling for more and less competitive markets and complete or incomplete price indexing (Treatments T1-T4), the SGG risk elicitation task was implemented with hypothetical rewards.

Table 4.2

Random effects GLS regression: Pricing explained by risk attitudes.

Dependent variable: price

Variable	Coefficient	Std. Errors
dummy_lose (t-1)	95.09 ***	5.63
period	-1.55 ***	0.18
dummy_t1	73.63 ***	18.54
dummy_t2	68.10 ***	18.59
dummy_t3	-4.57	18.64
pc1_scores	**-7.54** *	4.02
pc2_scores	**20.24** ***	6.95
constant	461.70 ***	14.53

Number of obs = 8820
Number of groups = 180

Breusch and Pagan LM test for random effects
chi2(1) = 13584.52
Prob > chi2 = 0.0000

(*) significant at 10% confidence level, (**) significant at 5% confidence level, (***) significant at 1% confidence level.

Following the estimates on Table 4.2 and abstracting from the specifics of the main experiment, we see that risk attitudes provide significant explanatory power for the pricing behavior observed. In fact, both the first and the second principal components are necessary to identify the effect of risk attitudes on pricing behavior. On one hand, the first component capturing safe choices is associated to more competitive pricing. That is, more risk-averse subjects set lower prices in order to avoid the risk of not having the lowest price indexed by the engine. On the other hand, the second principal component is associated with lower pricing. This means that subjects able to recognize the increased profitability of riskier choices across panels also realize that setting higher prices guarantees profits which do not depend on the excessive randomness of the search process.

4.4 Study 2: Exploring risk attitudes under large stakes and mixed outcomes

4.4.1 Tasks overview

As described in Study 1, the SSG lottery-panel task is designed to assess risk attitudes for small to moderate stakes in the domain of gains. However, it is well documented that in the context of financial and gambling situations individual risk attitudes are not only sensitive to the size of the stake but also to the decision domain (gains or losses). Markowitz's (1952) conjecture that risk aversion increases with stake size in the gains domain is widely supported by evidence from experimental studies using both real payoffs and hypothetical choices (e.g., Binswanger, 1980; Antoni Bosch-Domènech & Silvestre, 1999; Hogarth & Einhorn, 1990; Holt & Laury, 2002; Kachelmeier & Shehata, 1992; Kühberger, Schulte-Mecklenbeck, & Perner, 1999; B. J. Weber & Chapman, 2005; Wik, Aragie-Kebede, Bergland, & Holden, 2007), while evidence regarding the effect of the magnitude of stakes in risky behavior is not so clear in the case of losses (e.g., Etchart-Vincent, 2004; Fehr-Duda, Bruhin, Epper, & Schubert, 2010; Vieider, 2012).

Moreover, many real financial situations, such as investment, typically involve outcomes in the gains-losses domain (mixed outcomes). Risk attitudes for this particular type of prospects is found to be distinguishably different from risk attitudes in the pure gains or losses domain, where the general finding is that for moderate probabilities individuals behavior exhibits the reflection effect predicted by PT, that is, a tendency for risk-averse behavior in the domain of gains and risk seeking behavior in the domain of losses[31]. For mixed gambles, it is generally found, also as predicted by PT, and attributed mainly to loss aversion, that the majority of individuals behave risk averse. Also that they behave as more risk averse when the possible outcomes are mixed than when the outcomes are only gains (e.g., Brooks & Zank, 2005; Schoemaker, 1990; Wik et al., 2007)[32]. Furthermore, it has been documented that for mixed outcome lotteries risk aversion has a tendency to increase with the magnitude of the stake (e.g., Vieider,

[31] To be precise, the interaction between value function's shape (concave for gains - convex for losses) and small probabilities overweighting, leads to the "fourfold pattern of risk attitudes", that is, risk aversion for gains involving moderate probabilities and losses involving small probabilities and risk seeking for losses involving moderate probabilities and gains involving small probabilities.

2012; Wik et al., 2007). However, risky behavior in mixed lotteries presents more irregularities than choices in only gains lotteries. In mixed gambles some studies have also found risk seeking behavior in small stakes (e.g., Battalio, Kagel, & Jiranyakul, 1990; Kameda & Davis, 1990) or risk neutrality in hypothetical choices with high stakes and non-hypothetical small stakes (e.g., Ert & Erev, 2010). In order to assess these effects in risky behavior, we designed three additional treatments of the SGG lottery-panel task.

Recall that the original task consists of four different panels, every one of which contains ten different lotteries. In each lottery, subjects can win a prize (x) with a probability (p) and otherwise nothing. In all the new treatments we maintain the original winning probability range (from 1 to 0.1 in 0.1 steps) and the general four panel structure, that is, all the panels begin with a sure amount, which is increased as winning probabilities are decreased. These increments are larger as we move from panel 1 to panel 4.

In order to explore the effect of large stakes, we multiplied the original task prizes by 10.000, therefore in each lottery subjects can win a prize $(x{\cdot}10.000)$, with a probability (p) and otherwise nothing. We call this treatment Large Gains (Figure 4.7)

Panel 1

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	10,000	11,000	13,000	15,000	17,000	21,000	27,000	36,000	54,000	109,000
Choice										

Panel 2

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	10,000	12,000	15,000	19,000	23,000	30,000	40,000	57,000	90,000	190,000
Choice										

Panel 3

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	10,000	17,000	25,000	36,000	50,000	70,000	100,000	150,000	250,000	550,000
Choice										

Panel 4

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	10,000	22,000	38,000	57,000	83,000	120,000	175,000	267,000	450,000	1,000,000
Choice										

Figure 4.7 *The Large Gains treatment.*

To introduce losses on choices, we subtract 1 monetary unit to every payoff, therefore, in each lottery subjects can win a prize $(x-1)$ with a probability (p) and otherwise loose 1 monetary unit. We call this treatment Small Losses. Finally, we explore a combination of large stakes and losses. The payoffs in this treatment are obtained by multiplying the payoffs of the Small Losses treatment by 10.000, therefore subjects can win a prize $\left[(x-1)\cdot 10.000\right]$, with a probability (p) and otherwise loose 10.000 monetary units. We name this treatment Large Losses.

To summarize, we have the original SGG lottery task, henceforth Small Gains (SG) treatment, and the three new treatments: Large Gains (LG), Small Losses (SL) and Large Losses (LL). In both SL and LL treatments in each of the ten lotteries the possible outcome is either a loss or a gain. Therefore, these two treatments combine the domain of losses with the domain of gains (mixed gambles).

4.4.2 Implementation

A total of 170 subjects participated in this study. All of them were undergraduate students from the University of Granada, the sample includes 127 Economics and Business students and 43 Psychology students (62.52% females) aged on average 22.31 years (S.D.=2.73, age range=18-39).

The four treatments were administered in paper and pencil format and they were presented in the following order: SG-LG-SL-LL. Due to the nature and magnitude of the rewards choices in all treatments were hypothetical. Each participant completed four decisions in each of the four treatments. Consequently, we have a total of sixteen decisions per subject.

The first session took place on the 20th of April 2009 (N=127) and was conducted in the Faculty of Economics and Business in the framework of a course on behavioral economics. The second sample was collected at the University of Granada Experimental Physiology Laboratory between April and July of 2011. The Psychology students participated in individual sessions and received course credits for their participation. The augmented SGG task was part of the second session of a broader study that included several physiological measurements and questionnaires, not reported here.

Although there are some changes in conditions between the first and the second implementation, if we compare them we find no evidence to reject that the samples have been drawn from same distribution of choices for the first three treatments (K-S corrected p>0.10) while we reject that samples come from the same distribution for the

LL treatment (K-S corrected p<.01). Despite this small difference, we decided to jointly analyze the data of both implementations (N=170). We do not focus on comparing samples since we cannot disentangle the source of the difference. It could be either due to implementation conditions or to idiosyncratic difference between subjects.

4.4.3 Results

Intra-treatment comparison: To analyze the data, we start by focusing on a treatment comparison. Figure 4.8 shows the pooled panel choices for each of the four treatments, one in each quadrant. SG: Top-Left; SL: Bottom-Left; LG: Top-Right and; LL: Bottom-Right. In the Figure we can observe that a general feature of all the treatments is that a high percentage of choices show some degree of risk aversion (choices different than p=0.1 are for all treatments higher than 70%). Another pattern that emerges in the data is three frequency peaks, which correspond to the lotteries with winning probabilities: 1, 0.5 and 0.1. An additional characteristic of the data is a considerable increase in the frequency of sure ($p=1$) lottery choices compared to close to sure ($p=0.9$) reward lottery choices, which is consistent with probability distortions implied by PT or the so-called certainty effect.

In spite of these common features, it is clearly observable that for each treatment the frequency of choices follows a different distribution, a conclusion that is confirmed using a Kolmogorov-Smirnov test for each pair ($p<0.0001$). Regarding modal choices, in the SG treatment, the mode corresponds to $p=0.5$; in the SL treatment it is the risk neutral alternative (p=0.1). For the LL and LG treatments it is on the other extreme of the probability space, which corresponds to the safe choice $p=1$. Also, as expected, the most notorious distribution differences are observed when bigger changes are introduced, for instance, from SG to SL the distribution change is much less dramatic than between SG and LL.

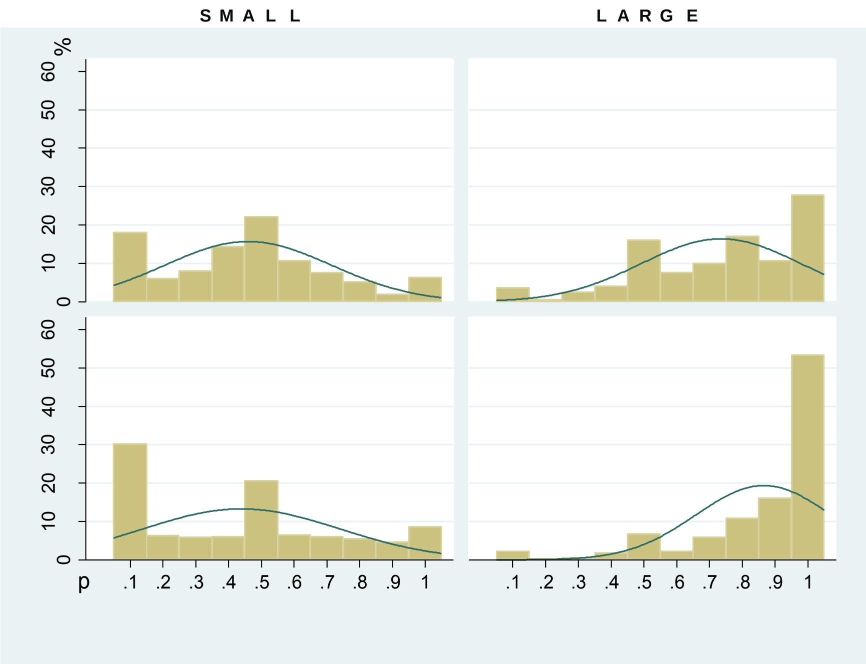

Figure 4.8 *Histograms of subjects' pooled probability choices across all panels, by treatment.*

From this we can draw some preliminary conclusions. At the aggregate level, we find that with respect to the effect of stake size, risk taking decreases with the magnitude of the stake in both gains and mixed lotteries, as is shown by a shift to the right in the distribution of choices. In particular, in the LG treatment, participants choose the sure payoff lottery with greater frequency (27.78% vs. 6.29% in SG) and a smaller number of risk neutral choices are made in this case (3.65% vs. 17.98% in SG). When high stakes are combined with the lotteries involving losses, the effect is even more dramatic (53.38% vs. 8.63% for p=1 and 2.21% vs. 30.12% for p=0.1). Therefore, for both gains and mixed (gains-losses) treatments, when stakes are increased the result is a shift to safer choices.

Also, at an aggregate level, we find an interaction effect between the size of the stake and the decision domain. In small stake treatments (SG and SL), risk taking is lower in the gains treatment (SG) than in the treatment that incorporates losses (SL). In particular, there is a considerable difference in the frequency of risk neutral choice (*p=0.1*) which is around 18% in SG treatment while in SL it is about 30%. On the other hand in large stake treatments (LG and LL), the contrary is observed, that is, risk taking decreases considerably in the treatment involving losses. In particular the modal choice *p=1* is observed with much higher frequency in the LL treatment (53.38% vs. 30.12%) and choices are, in general, concentrated near the safer lotteries.

If on the other hand we focus on within-subjects comparisons of pooled choices in each of the four treatments, we confirm our conclusion about the effect of stake size. There is a significant difference between choices in the SG and LG treatments and between choices in the SL and LL treatments (Wilcoxon rank-sum test p<0.001). The difference goes in the direction of lower degrees of risk taking in large stake treatments. At the same time within-subjects comparisons of choices between only gains and mixed lotteries reveal that at an individual level, there is a significantly lower degree of risk taking in the LG treatment compared to the LL treatment (Wilcoxon rank-sum test p<0.001). However, the difference is not so strongly significant between SG and SL (Wilcoxon rank-sum test p=0.0578).

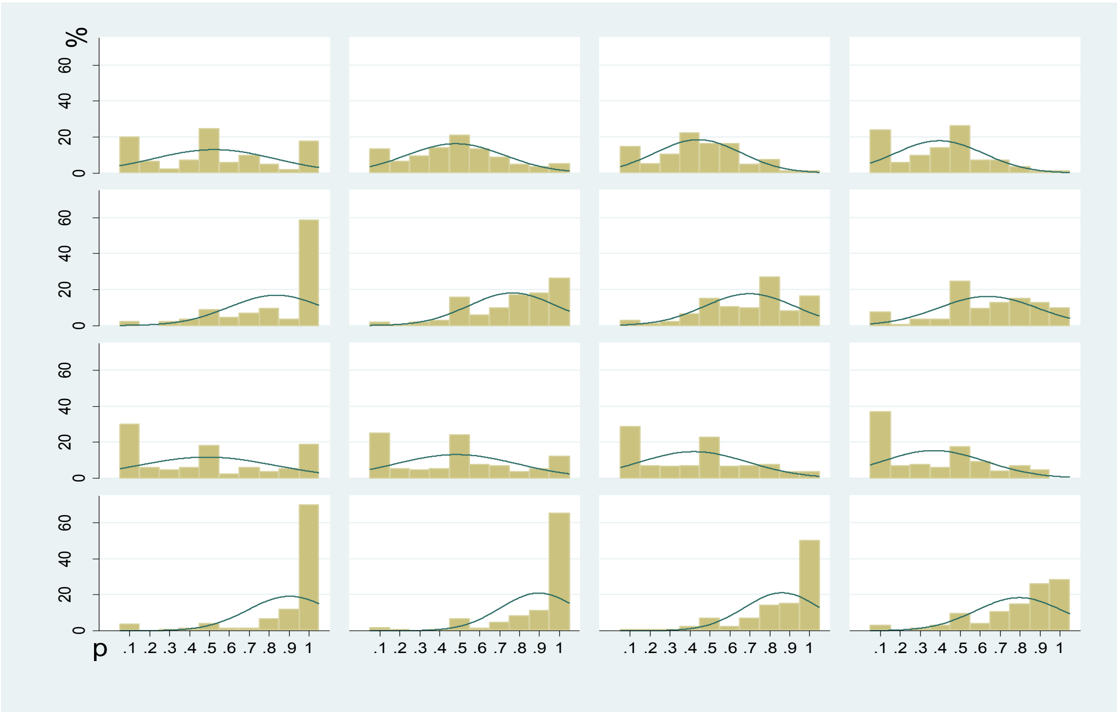

Figure 4.9 *Histograms of subjects' probability choices by panel and treatment.*

Intra panel comparison: To extend our analysis recall that, following the basic structure of the original SGG lottery task, each treatment contains four panels. Also, as we move from panel 1 to panel 4, the return to risk within each panel is larger. To go further into detail, we can analyze how the subjects' propensity to respond to different risk returns changes depending on the domain and the size of the stake. Figure 4.9 displays the sixteen histograms corresponding to the lottery choice frequencies in each of the treatments and panels. From left to right, we observe that choice frequencies of safer (riskier) lotteries decrease (increase). Therefore, at the aggregate level, subjects respond positively to the increase in return to risk. In general, the most significant response is observed between panels 3 and 4. These results are compatible with our previous results for the hypothetical implementation of the original SGG task (Study 1). Exploiting the within-subjects structure of our design, we find that in almost all the cases there is a significant choice variation across panels (Wilcoxon rank-sum test $p<0.001$, for 16 out of 18 panel pairs, exception are: panel 1=panel 2 in LL and panel 2=panel 3 in SG). This variation is positive in the sense that individuals respond positively to the increase in return to risk. On the other hand, the SL treatment is the one that shows less variation across panels both at an aggregated and individual (within-subjects) level. Actually, a significant shift is only observed between distant panels 1 and 4, and between panel 2 and 4, while for consecutive panels the difference is only significant across panel 2 and 3 ($p<0.01$).

Across treatments, on the other hand, we observe that for every panel (columns of Figure 4.9) an increase in the size of the stake decreases risk taking (SG vs. LG and SL vs. LL) while for changes between gains and mixed lotteries, an decrease in risk taking is only observed in the high stake treatments. The highest intensity of the change is in the three last panels. Therefore, the preliminary conclusions drawn from pooled panel choices are a stable feature of the treatments within panels. Moreover, comparing within-subjects choices between treatments for each particular panel, we find that for all panels risk taking is lower in large stake treatments than in the corresponding low stake treatment (Wilcoxon rank-sum test $p<0.001$, for all 8 pairs). Meanwhile the domain effect is only observable in the high stake treatments LG compared to LL where all 4 panel pairs are statistically different ($p<0.001$), while the within-subjects comparison between SG and SL reveals no statistical difference between any of the 4 panel pairs ($p>0.10$).

3.4.4 Principal Components Analysis

Like in Study 1, Principal Component Analysis (PCA) is suitable to reduce the number of dimensions obtained from the task maintaining the greater amount of variability. Also, performing this analysis we are able to check whether the choices maintain the same correlation structure across the different treatments. With this purpose, we first perform a separate PCA on the data obtained in each treatment.

Table 4.3

Principal Components Analysis

		Lottery				
		SG	LG	SL	LL	Pooled
Comp. 1						
	Panel 1	0.452 ***	0.467 ***	0.470 ***	0.487 ***	0.477 ***
	Panel 2	0.551 ***	0.533 ***	0.535 ***	0.532 ***	0.523 ***
	Panel 3	0.547 ***	0.539 ***	0.526 ***	0.540 ***	0.523 ***
	Panel 4	0.439 ***	0.455 ***	0.464 ***	0.434 ***	0.475 ***
Eigenvalue		2.654 ***	2.888 ***	2.962 ***	2.947 ***	3.242 ***
Percentage (%)		66.35	72.20	74.05	73.68	81.04
Comp. 2						
	Panel 1	0.625 ***	0.626 ***	0.638 ***	0.540 ***	0.646 ***
	Panel 2	0.263 ***	0.285 ***	0.296 ***	0.340 ***	0.275 ***
	Panel 3	-0.262 ***	-0.250 ***	-0.305 ***	-0.232 ***	-0.263 ***
	Panel 4	-0.668 ***	-0.681 ***	-0.642 ***	-0.734 ***	-0.662 ***
Eigenvalue		0.911 ***	0.726 ***	0.692 ***	0.770 ***	0.515 ***
Percentage (%)		22.78	18.15	17.30	19.25	12.87
Cumulative %		89.13	90.35	91.35	92.93	93.91

*** significant at 1% level of confidence.

As Table 4.3 shows, for all treatments, the first PC can be interpreted as a weighted average of the choices in the four panels and explains around 70% of the variance in each treatment. The second PC is, for every treatment, juxtaposing choices in panel 1 and 2 with choices in panel 3 and 4 The higher weight is for extreme panels 1 and 4 in all cases. The second PC for each treatment explains around 20% of the variance of the data, and together with the first PC it explains around 90% of the variance in the data. Therefore, the technique confirms the presence of two dimensions in our data: average

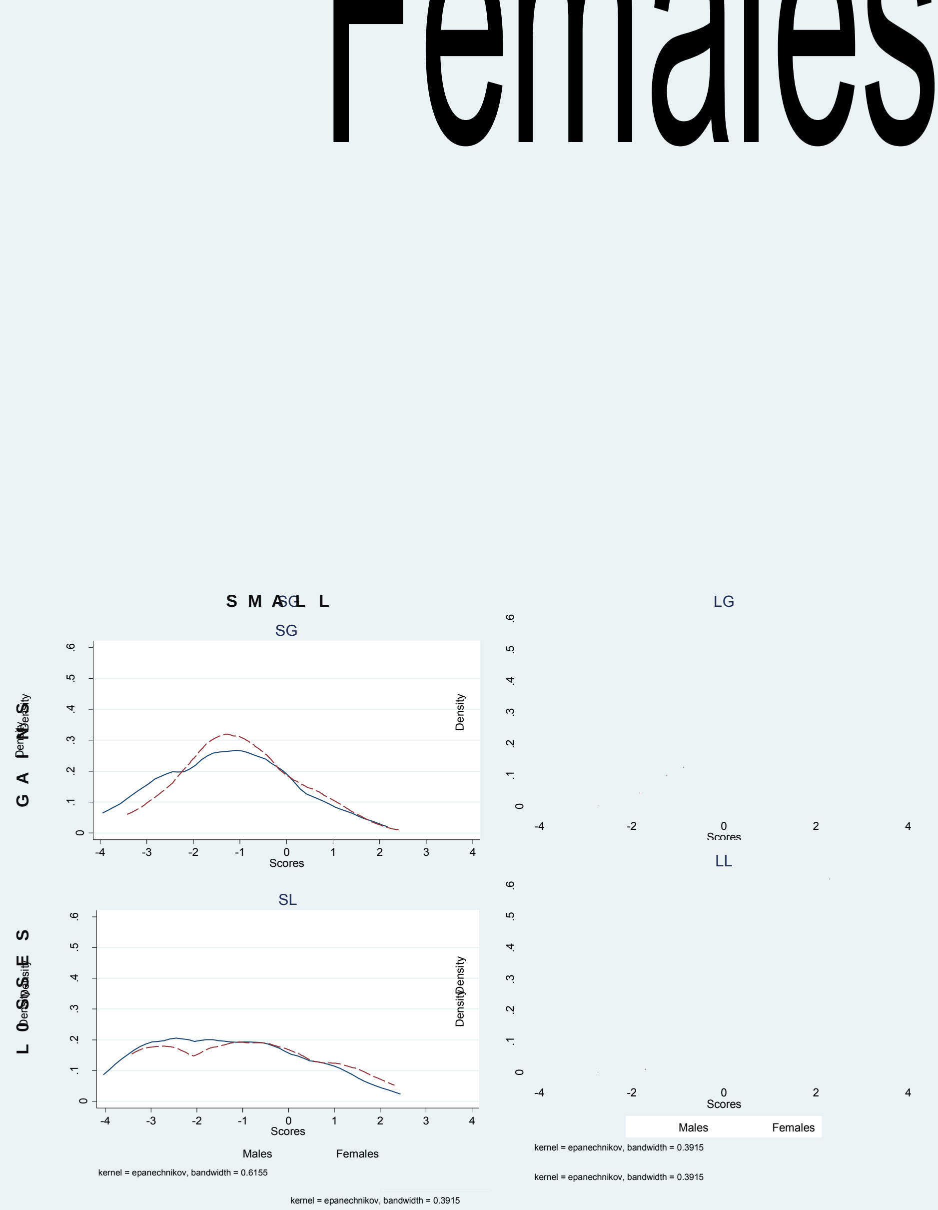

Figure 4.10 *Kernel density estimates for first component scores, by treatment and gender.*

From the representation of the first PC (risk taking) in Figure 4.10 we can tell for example, that the lower level of risk taking is found in the LL treatment and the higher in the SL treatment, and in general we reach the same conclusions as with the distribution of pooled data. Stake levels, both in the gains and the mixed domain decrease risk taking, and, there is an interaction effect between stake level and domain. For high stakes, we find the usual result that individuals exhibit a lower degree of risk taking in mixed outcome lotteries than in only gains lotteries, while in small stake lotteries, incorporating losses increases the level of risk taking. An additional feature which we observe is that for SG treatment the first PC has a close to normal distribution, slightly skewed to left (more risk taking) while in the SL treatment the distribution has a very low kurtosis, showing a high variance in the choices. This means that although at the aggregate level there is less risk taking in the SG than in the SL

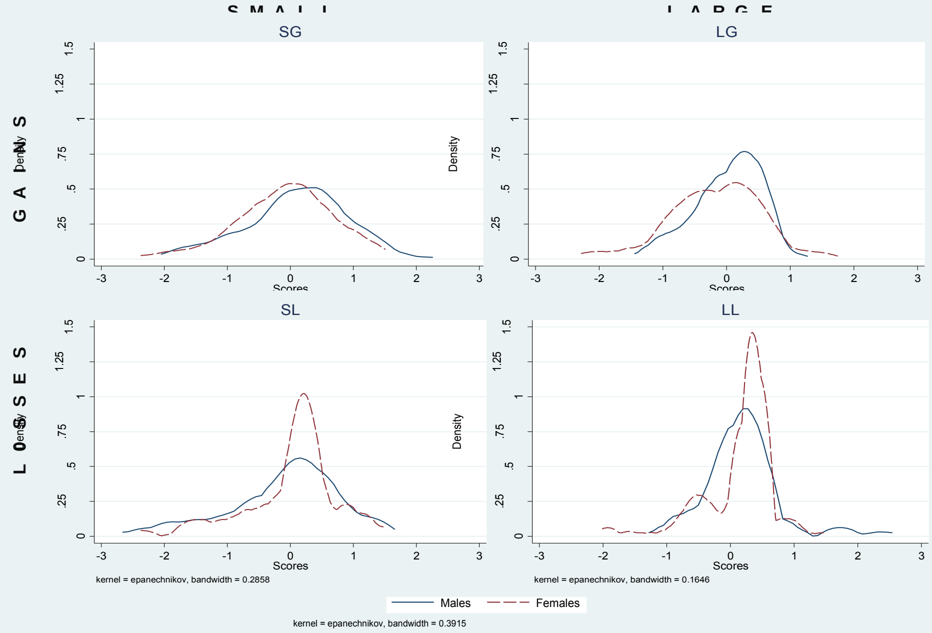

Figure 4.11 *Kernel density estimates for second component scores, by treatment and gender.*

On the other hand, using the second PC (Figure 4.11) we can extend the conclusions we obtained using the aggregate data. Considering both genders together, we find that compared to only gains treatments sensitivity to the return-to-risk premium is lower in the mixed lotteries treatments (SL and LL), as it is reflected by the important amount of values around zero. With respect to gender effects, it seems that in mixed outcome treatments, females tend to respond more positively to the increase in risk-return and also they behave more similarly to males. Males on the other hand show more of a positive sensitivity to return to risk in only gains treatments.

Another approach to analyze our data using PCA is to introduce all 16 choices together to reveal the correlation structure among panels and treatments. The result of this analysis is presented in Table 4.4.

Table 4.4

Principal Components Analysis, 16 choices

Component		Eigenvalue	Percentage (%)	Cumulative %
Comp. 1		5.280 ***	33.00	33.00
Comp. 2		3.119 ***	19.49	52.49
Comp. 3		1.744 ***	10.90	63.39
Comp. 4		1.505 ***	9.40	72.79
		Coefficients		
Treatment		**Comp. 1**	**Comp. 2**	**Comp. 3**
SG	Panel 1	0.210 ***	0.212 ***	-0.185
SG	Panel 2	0.240 ***	0.220 ***	-0.233
SG	Panel 3	0.289 ***	0.185 ***	-0.209
SG	Panel 4	0.250 ***	0.179 ***	-0.073
LG	Panel 1	0.255 ***	-0.197 ***	-0.284 ***
LG	Panel 2	0.261 ***	-0.237 ***	-0.350 ***
LG	Panel 3	0.304 ***	-0.185 ***	-0.303 **
LG	Panel 4	0.246 ***	-0.186 ***	-0.190
SL	Panel 1	0.220 ***	0.292 ***	0.242 ***
SL	Panel 2	0.245 ***	0.318 ***	0.272 **
SL	Panel 3	0.271 ***	0.303 ***	0.197
SL	Panel 4	0.242 ***	0.267 ***	0.103
LL	Panel 1	0.238 ***	-0.280 ***	0.231
LL	Panel 2	0.233 ***	-0.319 ***	0.290 *
LL	Panel 3	0.260 ***	-0.310 ***	0.334 ***
LL	Panel 4	0.217 ***	-0.236 ***	0.316 ***

* significant at 10% level of confidence, ** significant at 5% level of confidence,
*** significant at 1% level of confidence

Components eigenvalues and contribution percentages (top); Loads per component (bottom).

Again we find that the first PC can be interpreted as a weighted average of all 16 decisions and in this case explains one third of the variance of the data. Interestingly, the second component juxtaposes every panel from the small stake treatments (SG and SL) against every other panel corresponding to high stake treatments (LG and LL) and explains 19.5% of the variance of our data. The third PC juxtaposes every panel corresponding to treatments that only contain gains (SG and LG) against panels within the treatments that contain possible losses (SL and LL) and explains 10.9% of the variance of our data. Nevertheless, although the component is statistically significant, the contributions of some of the panels are not. The results of this exploratory technique confirm our previous observation that, in the context of our test, stake size creates a greater variability in choices than the domain.

4.5 Summary and conclusions

In the first study we proposed a simple task for the eliciting attitudes toward risky choice, the SGG lottery-panel task, which consists in a series of lotteries constructed to compensate riskier options with higher risk-return trade-offs. Using Principal Component Analysis technique, we show that the SGG lottery-panel task is capable of capturing two dimensions of individual risky decision making i.e. subjects' average risk taking and their sensitivity towards variations in risk-return. From the results of a large experimental dataset, we confirm that the task systematically captures a number of regularities such as: A tendency to risk averse behavior (only around 10% of choices are compatible with risk neutrality); An attraction to certain payoffs compared to low risk lotteries, compatible with over-(under-) weighting of small (large) probabilities predicted in PT and; Gender differences, i.e. males being consistently less risk averse than females but both genders being similarly responsive to the increases in risk-premium.

Another interesting result is that in hypothetical choices most individuals increase their risk taking responding to the increase in return to risk, as predicted by PT, while across panels with real rewards we see even more changes, but opposite to the expected pattern of riskier choices for higher risk-returns. Therefore, we conclude from our data that an "economic anomaly" emerges in the real reward choices opposite to the hypothetical choices. These findings are in line with Camerer's (1995) view that although in many domains, paid subjects probably do exert extra mental effort which improves their performance, choice over money gambles is not likely to be a domain in which effort will improve adherence to rational axioms (p. 635). Finally, we demonstrate that both dimensions of risk attitudes, average risk taking and sensitivity towards variations in the return to risk, are desirable not only to describe behavior under risk but also to explain behavior in other contexts, as illustrated by an example.

In the second study, we propose three additional treatments intended to elicit risk attitudes under high stakes and mixed outcome (gains and losses) lotteries. Using a dataset obtained from a hypothetical implementation of the tasks we show that the new treatments are able to capture both dimensions of risk attitudes. This new dataset allows us to describe several regularities, both at the aggregate and within-subjects level. We find that in every treatment over 70% of choices show some degree of risk aversion and only between 0.6% and 15.3% of individuals are consistently risk neutral within the same treatment. We also confirm the existence of gender differences in the degree of risk taking, that is, in all treatments females prefer safer lotteries compared to males. Regarding our second dimension of risk attitudes we observe, in all treatments, an increase in risk taking in response to risk premium increases.

Treatment comparisons reveal other regularities, such as a lower degree of risk taking in large stake treatments compared to low stake treatments and a lower degree of risk taking when losses are incorporated into the large stake lotteries. Results that are compatible with previous findings in the literature, for stake size effects (e.g., Binswanger, 1980; Antoni Bosch-Domènech & Silvestre, 1999; Hogarth & Einhorn, 1990; Holt & Laury, 2002; Kachelmeier & Shehata, 1992; Kühberger et al., 1999; B. J. Weber & Chapman, 2005; Wik et al., 2007) and domain effect (e.g., Brooks and Zank, 2005, Schoemaker, 1990, Wik et al., 2007). Whereas for small stake treatments, we find that the effect of incorporating losses into the outcomes is not so clear. At the aggregate level an increase in risk taking is observed, but also more dispersion in the choices, whilst at the within-subjects level the effect weakens. Finally, regarding responses to risk premium, we find that compared to only gains treatments sensitivity is lower in the mixed lotteries treatments (SL and LL). In general sensitivity to risk-return is more affected by the domain than the stake size.

After having described the properties of risk attitudes as captured by the SGG risk elicitation task and its three new versions, it is important to recall that the danger of using unidimensional descriptions of risk attitudes goes beyond the incompatibility with modern economic theories like PT, CPT etc., all of which call for tests with multiple degrees of freedom. Being faithful to this recommendation, the contribution of this essay is an empirically and endogenously determined bi-dimensional specification of risk attitudes, useful to describe behavior under uncertainty and to explain behavior in other contexts. Hopefully, this will contribute to create large datasets containing a multidimensional description of individual risk attitudes, while at the same time allowing for a robust context, compatible with present and even future more complex descriptions of human attitudes towards risk.

4.6 References

Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science*, *46*(11), 1497-1512.

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. *Management Science*, *53*(10), 1659-1674.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école americaine. *Econometrica*, *21*(4), 503-546.

Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, *76*(3), 583-618.

Aumann, R. J. (1987). Correlated Equilibrium as an Expression of Bayesian Rationality. *Econometrica*, *55*(1), 1–18.

Battalio, R. C., Kagel, J. H., & Jiranyakul, K. (1990). Testing Between Alternative Models of Choice Under Uncertainty: Some Initial Results. *Journal of Risk and Uncertainty*, *3*, 25-50.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7-15.

Berns, G. S., Capra, C. M., Chappelow, J., Moore, S., & Noussair, C. (2008). Nonlinear neurobiological probability weighting functions for aversive outcomes. *NeuroImage*, *39*, 2047-2057.

Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics*, *62*(3), 395–407.

Birnbaum, M. H., & Navarrete, J. B. (1998). Testing Descriptive Utility Theories: Violations of Stochastic Dominance and Cumulative Independence. *Journal of Risk and Uncertainty*, *17*, 49-78.

Bosch-Domènech, A., & Silvestre, J. (1999). Does risk aversion or attraction depend on income? An experiment. *Economics Letters*, *65*, 265–273.

Bosch-Domènech, A., & Silvestre, J. (2006). *Risk aversion and embedding bias* (Economics Working Papers 934). Barcelona, Spain: Department of Economics and Business, Universitat Pompeu Fabra. Retrieved from: http://www.econ.upf.edu/docs/papers/downloads/934.pdf

Brañas-Garza, P., Georgantzís, N., & Guillén, P. (2007). Direct and indirect effects of pathological gambling on risk attitudes. *Judgment and Decision Making, 2*(2), 126-136.

Brañas-Garza, P., Guillén, P., & López del Paso, R. (2008). Math skills and risk attitudes. *Economics Letters, 99*, 332-336.

Brooks, P., & Zank, H. (2005). Loss Averse Behavior. *The Journal of Risk and Uncertainty, 3*(2005), 301-325.

Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty, 2*, 95-104.

Camerer, C. F. (1995). Individual Decision Making. In J. Kagel & A. Roth (Eds.), *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.

Camerer, C. F., & Ho, T.-H. (1994). Violations of the Betweenness Axiom and Nonlinearity in Probability. *Journal of Risk and Uncertainty, 8*, 167-196.

Campo, S., Guerre, E., Perrigne, I., & Vuong, Q. (2011). Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders. *Review of Economics Studies, 78*, 112-147.

Carbone, E., & Hey, J. D. (2000). Which Error Story is Best? *Journal of Risk and Uncertainty, 20*(2), 161-176.

Cox, J. C., & Oaxaca, R. L. (1996). Is Bidding Behavior Consistent with Bidding Theory for Private Value Auctions? *Research in Experimental Economics, 6*, 131-148.

Etchart-Vincent, N. (2004). Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses. *Journal of Risk and Uncertainty, 28*(3), 217-235.

Fehr-Duda, H., Bruhin, A., Epper, T., & Schubert, R. (2010). Rationality on the rise: Why relative risk aversion increases with stake size. *Journal of Risk and Uncertainty, 40*(2), 147-180.

García-Gallego, A., Georgantzís, N., Navarro-Martínez, D., & Sabater-Grande, G. (2011). The stochastic component in choice and regression to the mean. *Theory and Decision, 71*(2), 251-267.

García-Gallego, A., Georgantzís, N., Pereira, P., & Pernías-Cerillo, J. (2005). *Competing against simulated equilibrium price dispersions: An experiment on Internet-assisted search markets* (NET Institute Working Paper 5-12). Retrieved from: http://archive.nyu.edu/bitstream/2451/28423/2/Georgantzis_05-12.pdf

Georgantzís, N., Genius, M., García-Gallego, A., & Sabater-Grande, G. (2003). *Testing for risk premium effects on individual choice under uncertainty* (Working Paper 43/03). Valencia, Spain: LINEEX, University of Valencia. Retrieved from: http://pascal.iseg.utl.pt/~cemapre/ime2002/main_page/papers/NikosGeorgantzis.pdf

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2002). Quantal response equilibrium. *Journal of Economic Theory*, *104*(1), 247–272.

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2003). Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, *45*(1), 97–113.

Goldstein, W. M., & Einhorn, H. J. (1987a). Expression theory and the preference reversal phenomena. *Psychological Review*, *94*(2), 236-254.

Goldstein, W. M., & Einhorn, H. J. (1987b). Expression theory and the preference reversal phenomena. *Psychological Review*, *94*(2), 236-254

Harrison, G. W., List, J. A., & Towe, C. (2007). Naturally occurring preferences and exogenous laboratory experiments: A case study of risk aversion. *Econometrica*, *75*(2), 433-458.

Harrison, G. W., Rutström, E. E., & Towe, C. (2009). Expected utility theory and prospect theory: one wedding and a decent funeral. *Experimental Economics*, *12*, 133-158.

Hey, J. D., & Orme, C. (1994). Investigating Generalizations of Expected Utility Theory Using Experimental Data. *Econometrica*, *62*(6), 1291-1326.

Hogarth, R. M., & Einhorn, H. J. (1990). Venture Theory: A Model of Decision Weights. *Management Science*, *36*(7), 780-803.

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

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

Hsu, M., Krajbich, I., Zhao, C., & Camerer, C. F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *The Journal of Neuroscience*, *29*(7), 2231-7.

Kachelmeier, S. J., & Shehata, M. (1992). Examining risk preferences under high monetary incentives: Experimental evidence from the People's Republic of China. *American Economic Review*, *82*(5), 1120–1141.

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

Kameda, T., & Davis, J. H. (1990). The Function of the Reference Point in Individual Risk Decision Making. *Organizational Behavior and Human Decision Processes*, *46*, 55-76.

Kühberger, A., Schulte-Mecklenbeck, M., & Perner, J. (1999). The Effects of Framing, Reflection, Probability, and Payoff on Risk Preference in Choice Tasks. *Organizational behavior and human decision processes, 78*(3), 204-231.

Lusk, J. L., & Coble, K. H. (2005). Risk perceptions, risk preferences and acceptance of risky food. *Agricultural Economics*, *87*(2), 393-405.

Prelec, D. (1998). The Probability Weighting Function. *Econometrica*, *66*(3), 497-527.

Rabin, M., & Thaler, R. H. (2001). Anomalies: Risk Aversion. *Journal of Economic Perspectives*, *15*(1), 219-232.

Rigas, G., Carling, E., & Brehmer, B. (2002). Reliability and validity of performance measures in microworlds. *Intelligence*, *30*(5), 463-480.

Sabater-Grande, G., & Georgantzís, N. (2002). Accounting for risk aversion in repeated prisoners ' dilemma games: an experimental test. *Journal of Economic Behavior & Organization*, *48*, 37-50.

Schoemaker, P. J. H. (1990). Are Risk-Attitudes Related across Domains and Response Modes? *Management Science*, *36*(12), 1451–1463.

Stott, H. P. (2006). Cumulative prospect theory 's functional menagerie. *Journal of Risk and Uncertainty*, *32*, 101-130.

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

Vieider, F. M. (2012). Moderate stake variations for risk and uncertainty, gains and losses: methodological implications for comparative studies. *Economics Letters*, Forthcoming.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Wakker, P., & Deneffe, D. (1996). Eliciting von Neumann-Morgen Utilities when Probabilities Are Distorted or Unknown. *Management Science*, *42*(8), 1131-1150.

Wang, M., Rieger, M. O., & Hens, T. (2010). How time preferences differ: Evidence from 45 countries (Research Paper 09-47). Switzerland: Swiss Finance Institute. Retrieved from: http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID1885349_code879164.pdf?abstractid=1481443&mirid=1

Weber, B. J., & Chapman, G. B. (2005). Playing for peanuts: Why is risk seeking more common for low-stakes gambles? *Organizational Behavior and Human Decision Processes*, *97*(1), 31-46.

Weber, E. U., & Hsee, C. (1998). Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes towards Perceived Risk. *Management Science*, *44*(9), 1205-1217.

Weber, E. U., & Hsee, C. (1999). Models and mosaics: Investigating cross-cultural differences in risk perception and risk preference. *Psychonomic Bulletin & Review*, *6*(4), 611-7.

Zuckerman, M. (1978). Sensation seeking. In H. London & J. Exner (Eds.), *Dimension of personality*. New York: John Wiley & Sons.

5. Appendix

5.1 Appendix Essay 1

5.1.1 Details on cognitive tasks implementation procedure

Raven's Standard Progressive Matrices Test: We implemented a computerized version using E-prime software (Schneider, Eschmann, & Zuccolotto, 2002) of the test originally proposed by Raven and Court (1948). In this test, participants were presented with sixty consecutive items, each one consisting of a matrix containing black-and-white abstract figures with one missing figure. For each matrix, response choices (figures) were presented. The participant's task was to choose among the response figures, the one that best completed the pattern. The sixty items were divided into five twelve items sets (A to E), with items within a set becoming increasingly difficult. Participants had as much time as they needed to complete this task. The total number of correctly answered items was used as the Raven's score.

Dot Counting Span Task: Following Pickering et al., (1999), a display booklet was placed in front of each participant consisting of pages each showing an area that contained either three, four, five or six stimuli (black dots). The participant was instructed to count aloud the number of dots and remember the count total for later recall. When the participant finished counting, the experimenter presented the next display. After a number of displays had been presented, the experimenter asked the participant to recall all the dot count totals in the order in which the displays were presented. The number of displays per series ranged from two to six, and they were presented in increasing order. Four series of each length were performed for a total of 20 series. The total number of correct responses was used as the dot-counting score.

N-Back Task: We implemented a computerized version of Kirchner (1958) using E-prime (Schneider, Eschmann, & Zuccolotto, 2002). In this task, participant was presented with a sequence of stimuli (phonologically distinct letters from the alphabet), and the task consisted of pressing a YES key when the letter in the screen matched the one from *N* steps earlier in the sequence and a NO key when it did not match the one from *N* steps earlier in the sequence. We presented three memory loads *N*, this is, 1-

Back, 2-Back and 3-Back. For each memory load *N* the participant performed a practice block (non scored sequence of 20 letters) to get familiar with the task, and two critical blocks (scored sequences of 30 letters each) in which there were 10 target and 20 no-target stimuli (20 target and 40 no-target stimuli per memory load). Individuals' scores obtained for each load consisted of the number of correct target responses (hits), incorrect target responses (misses), correct non target responses (right rejections), incorrect non target responses (false alarms) and missing responses, as well as reaction times for target, no-target and correct responses.

The Wisconsin Card Sorting Test: We used Heaton, Chelune, Talley, Kay, and Curtiss (1993) implementation of Grant and Berg (1948) task. In this task the experimenter presented four stimulus cards to the participant. The attributes of the cards were different in color (red, green, blue, or yellow), number (1, 2, 3 or 4) and shape (circle, cross, star or square). After observing the four cards, the participant was given a stack of additional cards and was then requested to put each card under one of four stimulus cards and to deduce the matching principle on the basis of feedback (correct, incorrect). We used twice each possible matching principle, with the following order: color, shape, number, color, shape, number. Each matching principle stayed the same until the participant correctly performed 10 consecutive matches, at which point the matching principle was changed (e.g., to shape). The task began and continued until either the participant had successfully achieved the 6 matching criteria or until the total number of target cards reached 128. The main dependent measure was the number of classical perseverative errors, which was the number of times participants failed to change matching criterion when the category changed and kept sorting the cards according to the previous, no longer correct matching principle.

The Stop-Signal Reaction-Time Task: We applied Verbruggen and Logan (2008) STOP-IT computerized task. In this task, participants were presented with a series of visual stimuli (squares or circles) and occasionally they heard a tone (stop-signal). Each participant had to perform a primary visual reaction time task, which was to press a key when a circle appeared in the screen and another when the square appeared; while occasionally a tone indicated them to stop their response to the primary task. Participants performed 3 blocks, in each block a sequence of 64 visual stimuli was presented, half of them circles and the other half squares, and a quarter of the times the visual stimulus was accompanied by a phonetic stop signal after a variable stop-signal delay. On both no-stop-signal trials and stop-signal trials, the stimulus remained on the screen until participants responded or until the maximal RT had elapsed. Participants were instructed to respond as quickly and accurately as possible to the go stimulus on no-stop-signal trials. The instructions emphasized that they should not slow down to wait for possible stop signals, since the software would detect this behavior and delay the appearance of the signal. The main result was whether or not participants withhold their response to the primary task when the stop-signal occurs. This is measured by the stop-signal reaction time, which is estimated from the signal response reaction time and the non-signal reaction time.

References

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, 38(4), 404-411.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sort Test Manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources, Inc.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55(4), 352-358.

Pickering, S. J., Baqués, J., & Gathercole, S. E. (1999). Bateria de tests de memòria de Treball. Barcelona: Laboratori de Memòria del U.A.B. (mimeo).

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002*). E-Prime: A user's guide.* Pittsburgh: Psychology Software Tools, Inc.

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in cognitive sciences*, 12(11), 418-24.

5.1.2 Instructions

The original instructions were in Spanish

Instructions for a Traveler's Dilemma (*R* = 5 treatment)

Welcome. Thank you for taking part and arriving punctually to this appointment. To accomplish this task you will be randomly paired with another person in this room.

Your unique task is to take a decision. Your earnings will depend on your decision and the decision of the other person with whom you are matched.

Along the whole session we are going to speak about coins of 5 cents of Euro. Therefore, a payment of 100 means that you receive: 100*5=500=5 Euros.

<u>The task:</u> You and the person you are paired with will each have to choose an entire number between 180 and 300 (including both).

<u>Rules:</u>

- If the two of you choose the same number, both will earn this amount of 5 cents coins.
- If the number chosen by each one is different, both of you will earn an amount equal to the minimum of both quantities (in coins of 5 cents).
    - In addition, the one of you who has chosen the highest number will be penalized with a "fine" of 5 coins.
    - These 5 coins will be received as prize by the one that has chosen the lowest number, this is, they will be added to his/her/your payment.

Summarizing, your earnings in this game will be (in coins of 5 cents of Euro):

- The number of your choice, if you and the other person chose the same number.
- The number of your choice PLUS (+) 5 coins if you have chosen the lowest number.
- The number chosen by the other person MINUS (-) 5 coins if you have chosen the highest number.

If you have any question, raise your hand and wait until we answer you privately. Do the calculation of your earnings departing from imaginary numbers for you and the other person. We will pass to verify that you have understood the structure of payments of this game.

USE THIS SPACE FOR NOTES and EXAMPLES (they will not affect your payments)

YOUR CHOICE: ____________________(An entire number between 180 and 300)

Instructions for a Traveler's Dilemma (*R* = 180 treatment)

Welcome. Thank you for taking part and arriving punctually to this appointment. To accomplish this task you will be randomly paired with another person in this room.

Your unique task is to take a decision. Your earnings will depend on your decision and the decision of the with whom you are matched.

Along the whole session we are going to speak about coins of 5 cents of Euro. Therefore, a payment of 100 means that you receive: 100*5=500=5 Euros.

<u>The task:</u> You and the person you are paired with will each have to choose an entire number between 180 and 300 (including both).

<u>Rules:</u>

- If the two of you choose the same number, both will earn this amount of 5 cents coins.
- If the number chosen by each one is different, both of you will earn an amount equal to the minimum of both quantities (in coins of 5 cents).
  - In addition, the one of you who has chosen the highest number will be penalized with a "fine" of 5 coins.
  - These 5 coins will be received as prize by the one that has chosen the lowest number, this is, they will be added to his/her/your payment.

Summarizing, your earnings in this game will be (in coins of 5 cents of Euro):

- The number of your choice, if you and the other person chose the same number.
- The number of your choice PLUS (+) 180 coins if you have chosen the lowest number.
- The number chosen by the other person MINUS (-) 180 coins if you have chosen the highest number.

If you have any question, raise your hand and wait until we answer you privately. Do the calculation of your earnings departing from imaginary numbers for you and the other person. We will pass to verify that you have understood the structure of payments of this game.

USE THIS SPACE FOR NOTES and EXAMPLES (they will not affect your payments)

YOUR CHOICE: ______________________(An entire number between 180 and 300)

General instruction for within-subjects part (Booklet)

In the following 6 sheets of paper, we you will present 6 different tasks.

Every sheet of paper corresponds to a task and in each one you will have to take a unique decision.

Once you finish taking your 6 decisions, please remain in your place. When everyone has finished we will pass gathering them.

To determine your earnings, you will be randomly paired with another person in this room.

At all time the anonymity of the two will be preserved.

The earnings of each one will be determined by your and him/her decisions. Only one of 6 tasks will be chosen randomly to be paid. That is to say, we will pay only for one of your six decisions, but neither you nor we know yet which of them will it be.

Remember that along the whole session a unit is equivalent to 5 cents of Euro. It means for example that 100 is equivalent to 100*5 cents = 500 cents = 5 Euros.

All payments will be carried out at the end, privately and preserving your anonymity.

5.2 Appendix for Essay 2

5.2.1 Full regression models

Table 5.1

Full Probit models Row player choices

Task	Variable	Dependent variable: Row chose *Top* (1)	(2)	(3)
Raven PM	RAVEN_P	**-0.019 *** (0.011)	**-0.018 *** (0.011)	**-0.018 *** (0.010)
Dot Counting Span	DCS_TOT	**-0.009** (0.121)	**-0.007** (0.121)	
N-back	2-BACK_SCORE	**-0.032** (0.073)		
	3-BACK_SCORE		**-0.008** (0.047)	
Wisconsin	WIS_PPERSM	**0.021 **** (0.010)	**0.021 **** (0.009)	**0.020 **** (0.009)
Stop-Signal	SS_RT	**0.015 **** (0.007)	**0.015 **** (0.007)	**0.014 **** (0.006)
	AGE	**-0.035** (0.138)	**-0.040** (0.138)	
	DUMMY_ROW	**0.519** (0.463)	**0.494** (0.459)	
	CONSTANT	**-1.053** (3.247)	**-1.325** (3.225)	**-1.758** (1.406)
Observations		81	81	81
Log likelihood		-21.011	-21.097	-21.797
χ^2 (n)		14.490	14.320	12.220
Prob > χ^2		0.043	0.046	0.005
McFadden R-squared		0.256	0.253	0.229

*10% confidence level, **5% confidence level, ***1% confidence level.
Standard errors in parentheses.
For X^2, n=7 for models 1 and 2 and. n=3 for model 3.

Models 1 and 2, include all variables, with the difference that model 1 includes the *N*-Back, N=2 and Model 3 is for *N*-Back, N=3 (both variables cannot be included at the same time because they are correlated). Model 3 includes only statistically significant explanatory variables.

Table 5.2

Full Probit models for Column choices

Task	**Variable**	Dependent variable: Column chose *Bottom*		
		(1)	(2)	(3)
Raven PM	RAVEN_P	**0.004** (0.006)	**0.004** (0.006)	
Dot Counting Span	DCS_TOT	**0.071** (0.075)	**0.078** (0.077)	
N-back	2-BACK_SCORE	**-0.030** (0.048)		
	3-BACK_SCORE		**-0.077 **** (0.036)	**-0.120 **** (0.059)
Wisconsin	WIS_PPERSM	**-0.002** (0.005)	**-0.002** (0.005)	
Stop-Signal	SS_RT	**-0.001** (0.003)	**0.000** (0.003)	
	AGE	**-0.025** (0.099)	**-0.043** (0.105)	
	DUMMY_ROW	**-0.498 **** (0.297)	**-0.513 *** (0.303)	**-0.931 *** (0.484)
	CONSTANT	**0.561** (2.401)	**1.203** (2.373)	**-1.758** (1.406)
Observations		81	81	81
Log likelihood		-52.327	-50.100	-51.067
χ^2 (n)		5.540	10.000	8.060
Prob > χ^2		0.594	0.189	0.018
McFadden R-squared		0.050	0.091	0.073

Standard errors in parentheses.
For X^2, n=7 for models 1 and 2 and. n=2 for model 3.

Models 1 and 2, include all variables, with the difference that model 1 includes the *N*-Back, N=2 and Model 3 is for *N*-Back, N=3 (both variables cannot be included at the same time because they are correlated). Model 3 includes only statistically significant explanatory variables.

5.2.2 Instructions

The original instructions were in Spanish

Instructions for an asymmetric matching pennies game (320, treatment)

To accomplish this task you will be randomly paired with another person in this room.
Your unique task is to take a decision. Your earnings will depend on your decision and the decision of the other person with whom you are matched.

The task: You and the person you are paired with, will each have to choose either A or B.

Your earnings in this game (in coins of 5 cents) will be:

- If you chose A and the other A, you earn 320 and he/she earns 40
- If you chose B and the other A, you earn 40 and he/she earns 80
- If you chose A and the other B, you earn 40 and he/she earns 80
- If you chose B and the other B, you earn 80 and he/she earns 40

If you have any question, raise your hand and wait until we answer you privately. Do the calculation of your earnings departing from imaginary numbers for you and another person. We will pass to verify that you have understood the structure of payments of this game.

USE THIS SPACE FOR NOTES AND EXAMPLES (they will not affect your payments)

YOUR CHOICE: ______________________ (A or B)

Instructions for an asymmetric matching pennies game

To accomplish this task you will be randomly paired with another person in this room.

Your unique task is to take a decision. Your earnings will depend on your decision and the decision of the other person with whom you are matched.

The task: You and the person you are paired with, will each have to choose either A or B.

Your earnings in this game (in coins of 5 cents) will be:

- If you chose A and the other A, you earn 40 and he/she earns 320
- If you chose B and the other A, you earn 80 and he/she earns 40
- If you chose A and the other B, you earn 80 and he/she earns 40
- If you chose B and the other B, you earn 40 and he/she earns 80

If you have any question, raise your hand and wait until we answer you privately. Do the calculation of your earnings departing from imaginary options for you and the other person. We will pass to verify that you have understood the structure of payments of this game.

USE THIS SPACE FOR NOTES AND EXAMPLES (they will not affect your payments)

YOUR CHOICE: ______________________ (A or B)

5.3 Appendix for Essay 3

5.3.1 The tasks

Instructions for the lottery-panel task (SG treatment)

Each one of the following "panels" shows you 10 lotteries. Each lottery has a probability of winning a prize. The prize is the amount of Euros shown below that probability. **If you do not win the lottery you earn 0€**. Remember that you have to choose one lottery in each one of the four panels. Mark with an X the space corresponding to your choice.

Panel 1

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	1.00	1.10	1.30	1.50	1.70	2.10	2.70	3.60	5.40	10.90
Choice										

Panel 2

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	1.00	1.20	1.50	1.90	2.30	3.00	4.00	5.70	9.00	19.00
Choice										

Panel 3

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	1.00	1.70	2.50	3.60	5.00	7.00	10.00	15.00	25.00	55.00
Choice										

Panel 4

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	1.00	2.20	3.80	5.70	8.30	12.00	17.50	26.70	45.00	100.00
Choice										

Instructions for the lottery-panel task (LG treatment)

Each one of the following "panels" shows you 10 lotteries. Each lottery has a probability of winning a prize. The prize is the amount of Euros shown below that probability. **If you do not win the lottery you lose 1€.** Remember that you have to choose one lottery in each one of the four panels. Mark with an X the space corresponding to your choice.

Panel 1

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	10,000	11,000	13,000	15,000	17,000	21,000	27,000	36,000	54,000	109,000
Choice										

Panel 2

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	10,000	12,000	15,000	19,000	23,000	30,000	40,000	57,000	90,000	190,000
Choice										

Panel 3

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	10,000	17,000	25,000	36,000	50,000	70,000	100,000	150,000	250,000	550,000
Choice										

Panel 4

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	10,000	22,000	38,000	57,000	83,000	120,000	175,000	267,000	450,000	1,000,000
Choice										

Instructions for the lottery-panel task (SL treatment)

Each one of the following "panels" shows you 10 lotteries. Each lottery has a probability of winning a prize. The prize is the amount of Euros shown below that probability. **If you do not win the lottery you earn 0€**. Remember that you have to choose one lottery in each one of the four panels. Mark with an X the space corresponding to your choice.

Panel 1

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	0.00	0.10	0.30	0.50	0.70	1.10	1.70	2.60	4.40	9.90
Choice										

Panel 2

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	0.00	0.20	0.50	0.90	1.30	2.00	3.00	4.70	8.00	18.00
Choice										

Panel 3

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	0.00	0.70	1.50	2.60	4.00	6.00	9.00	14.00	24.00	54.00
Choice										

Panel 4

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	0.00	1.20	2.80	4.70	7.30	11.00	16.50	25.70	44.00	99.00
Choice										

Instructions for the lottery-panel task (LL treatment)

Each one of the following “panels” shows you 10 lotteries. Each lottery has a probability of winning a prize. The prize is the amount of Euros shown below that probability. **If you do not win the lottery you lose 10,000€.** Remember that you have to choose one lottery in each one of the four panels. Mark with an X the space corresponding to your choice.

Panel 1

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	0.00	1,000	3,000	5,000	7,000	11,000	17,000	26,000	44,000	99,000
Choice										

Panel 2

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	0.00	2,000	5,000	9,000	13,000	20,000	30,000	47,000	80,000	180,000
Choice										

Panel 3

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	0.00	7,000	15,000	26,000	40,000	60,000	90,000	140,000	240,000	540,000
Choice										

Panel 4

Prob.	**1**	**0.9**	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
€	0.00	12,000	28,000	47,000	73,000	110,000	165,000	257,000	440,000	990,000
Choice										