

Rapport Final
Observatoire de l'Épargne Européenne / LaRGE

Choix de portefeuille, sur-confiance et biais de disposition

Appel d'offre de l'Observatoire de l'Épargne Européenne : Croyances, Connaissance, Information et Education Financière

Equipe Académique

Shaneera BOOLELL - GUNESH, Post-Doctorante, LaRGE, EM-Strasbourg, Université de Strasbourg & IAF, Université de Neuchâtel. E-mail: shaneera@unistra.fr

Marie-Hélène BROIHANNE, Professeur de Finance, LaRGE, EM-Strasbourg, Université de Strasbourg. E-mail: mhb@unistra.fr

Maxime MERLI (Porteur de Projet), Professeur de Finance, LaRGE, EM-Strasbourg, Université de Strasbourg. E-mail: merli@unistra.fr

Patrick ROGER, Professeur de Finance, LaRGE, EM-Strasbourg, Université de Strasbourg. E-mail: proger@unistra.fr

Avril 2010

RESUME FRANCAIS DES CONTRIBUTIONS

Le premier article *Disposition effect, investor sophistication and taxes: Some French Specificities*, (Boolell-Gunesh S., M. Merli et M.-H. Broihanne) publié dans la revue *Finance* en 2009 étudie l'effet de disposition associé aux investisseurs individuels sur le marché français. L'effet de disposition est un biais comportemental qui traduit la tendance à vendre plus aisément les titres en position gagnante que les titres en position perdante par rapport à leur prix d'achat. La présence de ce biais a été largement établie sur différentes bases de données de transactions d'investisseurs individuels à travers le monde. L'ensemble de ces travaux conclut, comme dans le travail original mené par Odean en 1998, que les investisseurs sujets à ce biais obtiennent en moyenne de faibles rentabilités.

Notre étude s'inscrit dans ce courant de recherche en établissant pour la première fois la présence de ce biais sur des données françaises. Le travail est mené sur une base de données rassemblant les transactions de 90 244 investisseurs individuels auprès d'une société de courtage en ligne sur la période 1999-2006. En nous appuyant sur la méthodologie de Odean (1998), nos résultats indiquent que les titres gagnants ont en moyenne 60% de chances de plus que les titres perdants d'être vendus par les investisseurs. En outre, nous montrons que les explications rationnelles de ces ventes (besoin de réallocation de portefeuille, coûts de transaction élevés pour les titres ayant un prix faible, etc.) ne peuvent expliquer le phénomène observé.

Cette tendance des investisseurs à réaliser leurs gains et à conserver les titres perdants en portefeuille peut être atténuée par la sophistication des intervenants. Nous considérons qu'un investisseur « sophistiqué » est familier des marchés financiers ce qui le rend moins susceptible d'être sujet au biais de disposition. Sur notre base de données, le degré de sophistication est approché par le recours aux produits dérivés (warrants), la vente à découvert et la diversification des portefeuilles à l'international. Le biais de disposition n'est pas éliminé pour les investisseurs identifiés comme « sophistiqués » selon les critères de notre base de données.

Enfin, contrairement aux résultats des autres études, le biais de disposition n'est pas systématiquement inférieur au mois de décembre par rapport au reste de l'année. En effet, la

fin de l'année fiscale incite habituellement les investisseurs à vendre les titres perdants pour diminuer le montant global des plus-values imposables. Sur notre base de données, l'absence de ce phénomène fiscal soulève de nouvelles questions spécifiques à l'organisation de l'investissement individuel. En France, les individus ont la possibilité d'ouvrir un ou plusieurs compte-titres ainsi qu'un Plan Epargne en Actions, ces deux types de support permettant de manière équivalente la passation d'ordres financiers. Le compte PEA offre en outre un régime fiscal avantageux qui permet la non imposition des plus-values réalisées au-delà des cinq premières années d'existence du compte. Nous étudions alors l'impact fiscal de l'utilisation de ces comptes sur l'ampleur de l'effet de disposition et montrons qu'en moyenne celui-ci est inexistant.

Sur la même base de données, le second article, *Trading activity and Overconfidence: First Evidence from a large European Database*, (Boolell-Gunesh S., M. Merli) étudie la présence de surconfiance parmi les investisseurs individuels français. Trois méthodes visant à évaluer la rentabilité des stratégies d'investissement retenues par les investisseurs sont mises en œuvre et nos résultats montrent que quelle que soit la méthodologie considérée, les investisseurs individuels sont surconfiants et échangent trop fréquemment. Les stratégies adoptées par ceux-ci sont en moyenne sous-optimales, les titres achetés sous-performant systématiquement les titres vendus sur différents horizons d'investissement. Ces résultats confirment les résultats obtenus dans d'autres pays, en particulier les Etats-Unis (Odean (1999), Barber and Odean (2001), Barber and Odean (2002) and Glaser et Weber (2007)).

Enfin, dans la longue liste des énigmes ayant intrigué les économistes et financiers depuis plus de 50 ans figurent en bonne place l'acquisition simultanée de contrats d'assurance et de billets de loterie ainsi que la sous-diversification « chronique » des portefeuilles des investisseurs individuels. Dans l'article, *Testing alternative theories of financial decision making: a survey study with lottery bonds*, (P. Roger), nous groupons ici ces deux « puzzles » car les solutions proposées dans la littérature depuis l'origine reposent sur les mêmes idées qu'on peut synthétiser en trois points :

- 1) Les fonctions d'utilité peuvent être convexes à certains niveaux de richesse.
- 2) Les distributions de probabilités des rentabilités ne sont pas symétriques et les individus ont une préférence pour un moment d'ordre 3 positif.
- 3) L'évaluation des perspectives risquées réalisées par les agents économiques ne s'appuient pas sur les probabilités réelles. Ils opèrent une distorsion qui favorise les

rentabilités extrêmes, positives et/ou négatives selon les théories au détriment des rentabilités moyennes ou faibles.

Nous avons cherché à déterminer quels étaient les critères de choix des individus lorsqu'ils font face à des distributions de probabilité dont la skewness est positive et plus précisément, nous cherchons à savoir s'ils maximisent, selon la théorie standard, l'espérance d'une fonction d'utilité croissante et concave, ou s'ils obéissent à un modèle alternatif comme la théorie des perspectives de Tversky et Kahnemann (1992) ou à celle des anticipations optimales (Brunnermeier et Parker, 2005, Brunnermeier et al., 2007).

Pour ce faire, nous avons administré un questionnaire à 337 étudiants dans quatre universités et Business Schools différentes en proposant des choix d'obligations à lots dont les rendements présentent, par construction, une skewness positive.

Les résultats obtenus mettent en évidence de manière très nette un comportement qui s'accorde à la théorie des anticipations optimales. Plus de 56% des répondants ont des choix conformes à cette approche alors qu'un peu plus de 20% des choix sont compatibles avec l'espérance d'utilité.

Ces résultats montrent d'une part l'attractivité d'un investissement à skewness positive mais aussi l'attraction pour l'actif présentant le « meilleur résultat possible » le plus élevé. Ils permettent d'expliquer pourquoi les Premium bonds britanniques sont aussi populaires et incitent à la réflexion en matière d'incitation à l'épargne. Des comptes d'épargne associés à des loteries sont proposés dans de nombreux pays (Guillen et Tschöegl, 2002) et incitent à l'épargne les ménages à faible revenu.

Are French Individual Investors reluctant to realize their losses?

Boolell-Gunesh. S¹ – Broihanne M-H². – Merli M.

January 2009

Abstract

We investigate the presence of the disposition effect for 90 244 individual investors using a unique large brokerage account database between 1999 and 2006. Our main results show that individual investors demonstrate a strong preference for realizing their winning stocks rather than their losing ones. However, the fiscal impact in France appears to be moderate relative to the one observed in other countries. Taking French specificities such as, the way short sales are realized and the existence of tax free account (PEA account) into account, show that: a) the behavioral bias is not eliminated for sophisticated individual investors; b) the change of “tax account type” does not imply any change in investors’ behavior

JEL Classification : G 10

Résumé

Nous étudions la présence d’un effet de disposition pour 90 244 investisseurs individuels français à partir d’une base de données de transactions individuelles sur la période 1999-2006. Les principaux résultats montrent que ces investisseurs ont une préférence marquée pour la réalisation de leurs gains plutôt que de leurs pertes. L’impact fiscal semble en France plus modéré que dans d’autres pays

En outre, l’existence de possibilités de ventes à découvert (SRD) et de comptes PEA en France permet de mettre en lumière de façon originale a) la persistance de ce biais pour des investisseurs sophistiqués et b) le faible impact du régime fiscal sur le comportement des investisseurs.

JEL Classification : G 10

¹Corresponding author : LaRGE / EM Strasbourg Business School, Strasbourg University, 61 avenue de la forêt noire, 67085 Strasbourg Cedex, France. E-mail : shaneera@cournot.u-strasbg.fr

² LaRGE / EM Strasbourg Business School, Strasbourg University E-mail: mhb@cournot.u-strasbg.fr / merli@cournot.u-strasbg.fr.

I-Introduction

Recent research in behavioral finance has demonstrated that investment behavior is not always consistent with the assumptions of perfect rationality generally made in the field. More precisely, this behavior has sometimes been shown to be systematically different from what is implied by normative models of standard finance theory.

One of the most widely documented behavioral biases is the disposition effect. This effect describes the tendency, at any given point in time, to more readily sell winners than losers, winners and losers referring to assets that have appreciated or depreciated since purchase. In this framework, researchers have shown that investors who are prone to the bias earn poor subsequent returns on their portfolio (Odean, 1998). Of course, rational reasons can justify this behavior: portfolio rebalancing or higher trading costs of low priced assets, for instance. However, none of these reasons has been found convincing enough by researchers.

Starting with Shefrin and Statman (1985), a number of researchers among others have documented the effect: Lakonishok and Smidt (1986) on aggregate volumes, Odean (1998), Shapira and Venezia (2001), Dhar and Zhu (2006) on individual data³.

However, if the presence of the disposition effect has been answered in some countries, no such research has yet been carried out in France.

Our paper fits this loophole by investigating the trading records of 90 244 individual investors at a French discount brokerage house between 1999 and 2006. As a result, the first contribution of this study is to be the first one on the French market and the most comprehensive in the European context⁴. We find that investors show a strong preference for realizing paper gains rather than their paper losses and that this behavior cannot be explained, for instance, by a desire to rebalance portfolios.

³ Note that Weber and Camerer (1998) and Weber and Welfens (2006) bring experimental evidence of this bias.

⁴ The only European research dealing with the disposition effect on individual data concerns 3 079 accounts (Weber and Welfens, 2006).

We expect some investors to be more sophisticated than others. To be precise, investors are ranked as sophisticated ones if they trade derivative assets, internationally diversify their portfolio or use short selling facilities (French SRD). Based on this original approach, our second contribution demonstrates that sophisticated traders are also subject to the bias which leads us to conclude that sophistication attenuates but does not eliminate the disposition effect.

At the aggregate level, we show that the impact of the tax year effect is clearly less important in France than in other countries. French specificities i.e the existence of PEA account (Plan d'Epargne en Actions) give a unique opportunity to investigate more deeply the global impact of tax on the selling behavior of investors on the financial market. Actually, these accounts offer an interesting tax framework to their holders in the sense that capital gains are tax free if the account has been kept for more than 5 years. In this framework, we study the disposition effect for holders of PEA accounts before and after the end of the 5-year period. We show that individual investors do not seem to change their investment behavior according to the type of fiscal account held (PEA or traditional). This original is our third important contribution.

The rest of the article is organized as follows. The second section presents an overview of previous research on the disposition effect. Section III describes the data and introduces the methodology. Section IV is dedicated to the description of our main results and comments and finally we conclude in section V.

II-The Disposition Effect

The disposition effect is the tendency of investors to hold losers (losing stocks) too long and sell winners (winning stocks) too soon. This phenomenon was first documented by Shefrin and Statman (1985) in a study of mutual fund performance. Subsequent papers based on market data (Lakonishok and Smidt, 1986; Ferris, Haugen and Makhija, 1988) showed that volume for winning stocks on the NYSE and the Amex exceeds that for losers.

From a theoretical point of view⁵, many explanations of the disposition effect have been proposed in the literature. The most common explanation is based on the assumption of prospect theory preferences (Tversky and Kahneman, 1979, Kahneman and Tversky 1992) and, more precisely, on the S-shaped valuation function assumed in this model. According to this theory, investors evaluate gains and losses with respect to a reference point; the buying price is the most commonly used reference point. When a stock price is higher than the buying price (or more generally than the reference price), the investor is in the concave part of his valuation function and is hence risk averse. He may sell the stock if the expected return is perceived as too low. After a price drop, the investor is in the convex part and keeps the stock because he has become risk seeking. Following Shefrin and Statman (1985), some authors have used this argument to justify the existence of disposition investors (Odean (1998) and Weber and Camerer (1998), for example). In other words, when agents are risk-averse over gains and risk lovers over losses, they prefer to realize paper gains and to keep paper losses⁶. A second explanation is based on an irrational belief in mean reversion of stock prices, which states that investors believe poorer-performing stocks will rebound, and that better-performing stocks will decline in price. Briefly speaking, after a price increase, the investor believes that the probability of a price drop in the next period is higher than the one of another price increase (Shu et al. (2005), Weber and Camerer (1998)).

A third group of explanations argue that the disposition effect may be due to the desire to rebalance portfolios or to avoid higher transactions costs on low-priced assets. However, it has been shown in many studies that when controlling for rebalancing and share prices, the disposition effect is still observed and that the investments the investors choose to sell continue in subsequent months to outperform the losers they keep (see Odean (1998), Brown *et al.* (2006), for example).

A last explanation of the disposition effect is proposed by psychologists who work on the theory of entrapment or escalation of commitment (Staw (1979), Brockner (1992)). In an investment context, the question is to know if it is better to keep a losing investment, to

⁵ For experimental studies of this bias, see for example, Weber and Camerer (1998), Chui (2001), Weber and Welfens (2006) and Rubaltelli *et al.* (2005).

⁶ Barberis and Xiong (2006) (see also Hens and Vlcek (2005)) show that the disposition effect is observed for some values of the expected stock return and the horizon of the investor, but they also find the opposite effect for other reasonable values of these parameters. Note that Barberis and Xiong (2008) give some new theoretical explanation of the disposition effect based on a “realization utility”.

increase the stake (to break even), or to sell the losers and choose other stocks to invest in (Zuchel, 2001).

Finally, the disposition effect can also refer to preferences, including the idea that investors seek pride and want to avoid regret when choosing investment (Shefrin and Statman (1985)). This interpretation has recently been developed by Muermann and Volkman (2006). The authors argue that loss aversion alone cannot explain the disposition effect as shown by Barberis and Xiong (2006) and Hens and Vlcek (2005) and they include the anticipation of regret and pride in a dynamic portfolio choice setting⁷.

From an empirical point of view, the disposition effect is now well documented on individual data. Odean (1998) was the first to study the decision process of individuals on an important database of 10 000 accounts with a total of 97 483 transactions between 1987 and 1993.

He found that the proportion of realized gains is significantly higher than the proportion of realized losses (except in December), giving evidence of a disposition effect in individual investors' behavior.

Later studies on individual data gave rise to similar results for the behavior of employees (Heath *et al.*, 1999), and for stocks in other countries than the US (Shapira and Venezia, 2001, for Israel, Grinblatt and Keloharju, 2001, for Finland, Chen *et al.*, 2004, for China, and Shu *et al.* (2005), for Taiwan, Brown *et al.*, 2006, for Australia). The disposition effect also appears to be positive on average but of different magnitude across countries and across investors. For example, Barber *et al.* (2007) show that Taiwanese investors are much more reluctant to realize their losses than U.S investors. They interpret their findings by the fact that Taiwanese traders exhibit a stronger belief in mean reversion than U.S traders.

Note that at the individual level, the disposition effect could vary across individual investors. Concerning this level of analysis, Dhar and Zhu (2006) confirm the presence of a significant disposition effect on average but show that one-fifth of the investors exhibit the opposite behavior and that the disposition effect is stronger for less sophisticated investors. Finally, the disposition effect is also detected in the investment decisions of professional traders. (Shapira and Venezia (2001), Genesove and Mayer (2001), and Barber *et al.* (2007), for example)⁸.

⁷ For experimental evidence, see O'Curry, Fogel and Berry (2006).

⁸ Coval and Shumway (2005), Frino *et al.* (2005) and Locke and Mann (2003) obtain the same kind of results on different futures markets.

The next section presents the original and proprietary dataset over which we analyze the disposition effect.

III- Data and empirical design

The anonymous data for this study comes from Cortal Consors, a large French discount brokerage house. We obtained transaction data for all active⁹ accounts over the period 1999-2006, that is a total of 9 619 898 transactions, with 5 074 732 buy orders and 4 545 166 sell orders, for 92 603 investors. Data are contained in three files: trades, investors and fees. The trades file combines the following information for each trade: ISIN code of the asset, type of asset (common stocks, bonds, certificates, warrants), buy-sell indicator, sell short indicator, date, quantity and amount in Euros, place of quotation, account type (taxable versus tax-free account or French “PEA”), media used to place the order, order type. In the investors file, some demographical characteristics of investors are gathered: date of birth, sex, date of entry in and exit of the database, opening and/or closing dates of all accounts, place of living, and yearly number of trades. Finally, the fees file contains monthly fees paid by each investor and indicates whether they are trade fees or short sales fees.

In order to study the disposition effect, we extracted a dataset that only includes trades for common stocks. This dataset contains 8 464 518 trades, with 4 447 678 buy orders and 4 016 840 sell orders, made by 90 244 investors over 4 377 assets. For each stock, we build a file containing historical daily prices over the period 1999-2006. In this respect, securities ISIN codes are used to collect price data and information on splits and dividends through Fininfo¹⁰, the French data provider. At this step, some trades (less than 1%) were deleted from the dataset because, either we did not find data corresponding to the ISIN code (534 codes out of 4377 ones). The final database, for which all prices are available, gathers 8 438 885 trades (4 426 894 purchases and 4 011 991 sales) for 90 079 investors.

In the context of French markets, three points should be outlined. First, individual investors may trade shares on a tax-free account, called PEA (Plan d’Epargne en Actions), or, as in

⁹Over the period 1999-2005, active accounts are those with at least one transaction over 2 years (consecutive or not). For the last year of the sample, accounts are active if they hold at least one transaction over the entire year.

¹⁰www.Fininfo.com

other countries, on a traditional asset account. PEA accounts are very popular because banks mainly distribute these accounts to their customers as a first experience with trading on the stock market. Moreover, and more importantly for the scope of this paper, PEA accounts allow to realize tax-free capital gains if the account was opened at least 5 years ago. In our dataset, 10911 investors hold only PEA accounts and 35 598 investors hold both PEA and traditional accounts.

The second point relies on the international diversification of investors' trades. Only 9,3% of trades deal with non-French shares. It is not a surprising result because of the well-documented home-bias (Huberman, 2001). Figure 1 provides precisely the distribution of transactions across regions on our dataset. At the individual level, 54% of investors realize at least one trade on these foreign assets and we call them "international traders". Third, French individual investors have a very easy access to short sales and 1095 investors realize all trades using SRD orders; we call them "SRD investors"¹¹.

Figure 1 about here

The typical investor is a male (86,42%) and is 42 years old on average. Table 1 provides descriptive statistics on trading behavior of investors. The average number of assets per trade is nearly¹² 460. During the period 1999-2006, investors realized more than 90 trades amounting to an average of more than 3 800€ per trade (3 696€ for buy and 4 011€ for sell). As the median trade size, number and amount are respectively 60 assets, 22 trades and roughly 2000 €, we conclude that there is a considerable heterogeneity in the trading behavior of investors. On average, investors are active half of the time (4 years over the 8).

Table 1 about here

¹¹Note that SRD ("Système à Règlement Différé") is a French market specificity which allows investors to leverage and short sell.

¹² This relatively high number of stocks per trade is mainly due to some huge trades on penny stocks.

In this paper, we use the methodology given in Odean (1998) to compute the disposition effect and the following example provides an explicit example. Suppose that two investors (I and II) are active and that 3 stocks (X, Y, Z) are quoted on the financial market. Table 2 summarizes all investor's trades on the whole period of study (one month for this example). This table contains the dates (first column) when at least one of the two investors takes a position (Buy or Sell). "Asset" and "Price" indicate respectively the number of assets and the average price of the asset during the day. Columns "type" give the investor trade indicator: purchase (B for Buy), sale (S for sell) or nothing (NO). "Nber" indicates the trade volume.

Table 2 about here

Each day an investor sells securities, we determine whether the security is sold for a gain or a loss by comparing its selling price to its average purchase price. When the position changes and stocks are bought, the average purchase price is then adjusted.

Therefore, each sale is counted as a realized gain (RG) or a realized loss (RL). Each stock in the portfolio at the beginning of each day that is not sold during that day is considered to be an unrealized (paper) gain or loss. Paper gains or losses are defined by comparing the high and low daily price of the stock to its average purchase price. If these daily prices are above their average purchase price, the trade is counted as a paper loss (PL); in the opposite case, it is counted as a paper gain (PG); otherwise, neither a paper gain nor a loss is accounted for. All gains and losses are calculated after adjusting for splits. Following Odean, we choose the reference price to be the average purchase price.

To illustrate this methodology, in the example of table 2, on October the 10th, prices of X and Y are higher than their average purchase prices (contrary to the price of Z) and the first investor chooses to sell 10 stocks X and to keep his position on Y and Z.

Then, on this date and for this investor, we compute 1 realized gain (stocks X sold), 0 realized loss, 1 paper gain (stocks Y) and 1 paper loss (stocks Z). Table 3 summarizes for the two investors, the values of RG , RL , PG , PL , for all selling days.

Table 3 about here

It is important to notice that the four estimates, RG , RL , PG and PL could obviously not be computed for portfolios containing only purchases or sales, portfolios with only one trade or only one asset traded and for sales for which no previous purchase was identified. In this article, the final number of trades for which the preceding methodology can be applied is 8 230 826.

The last step of the methodology consists of using these key values (RG , RL , PG and PL) to compute the proportion of realized gains (PGR) and the proportion of realized losses (PLR) according to the following rules:

$$PGR = \frac{N_{RG}}{N_{RG} + N_{PG}}$$

$$PLR = \frac{N_{RL}}{N_{RL} + N_{PL}}$$

$$DE = PGR - PLR$$

where N_{RG} , N_{PG} , N_{RL} , N_{PL} denote the number of realized gains, the number of potential gains (paper gains), the number of realized losses and the number of potential losses (paper losses).

In this paper, the measure of the disposition effect is defined as the difference $DE = PGR - PLR$. When this difference takes a positive value, it indicates that investors are more prone to realize gains than losses. In our example, the last row of Table 3 (TOTAL) gives $N_{RG} = 3$, $N_{RL} = 2$, $N_{PG} = 1$ and $N_{PL} = 3$. Finally, $PGR = 0.75$, $PLR = 0.4$ and $DE = 0.35$.

It is important to notice that these values are computed across investors assuming that each sale for a gain (or a loss) and paper gain (or paper loss) on the day of the sale are separate independent observations. In this context, we test the following hypothesis:

H_0 : *Proportion of Gains Realized* \leq *Proportion of Losses Realized*

The Z-statistic (distributed normally) is applied to test this hypothesis where:

$$Z = \frac{PGR - PLR}{\sqrt{\frac{PGR(1 - PGR)}{N_{RG} + N_{PG}} + \frac{PLR(1 - PLR)}{N_{RL} + N_{PL}}}}$$

Note that assuming the independence at an account level (instead of at a transaction level) PGR , PLR and DE could be measured for each investor (instead of at an aggregate level)¹³. The global disposition effect is then defined as the average account disposition effect. In our example, Total I and Total II give the values of N_{RG} , N_{PG} , N_{RL} , N_{PL} that are used to compute the disposition effect at an individual level. After basic calculations the value of the average disposition effect is 0.415 (0.33 for the first investor and 0.5 for the second)¹⁴.

This simple illustration shows that the two measures of ED give obviously different results and even if at an aggregated level investors seem to suffer from the disposition effect, the disparity between investors may be very important.

In the following sections, the disposition effect is first studied globally based on the assumption on independence at the transaction level. Then we study the presence of the disposition effect among sub-groups of traders. In the last section, we measure the impact of the tax account type on the behavior of investors and then use an individual measure of the disposition effect.

IV – General results and discussion

IV – 1 Disposition effect and sophistication

In this section, we present the results at the aggregate level. Based on 4 011 991 sales, we compute a total of 1 998 924 disposition effects for 57 153 investors¹⁵. For sake of simplicity, investors for whom a disposition effect is computed are called “investors” in the rest of the paper. We study the aggregate disposition effect (see tables 2 and 3 for an example) by considering that each sale that results in a realized or paper gain/loss constitutes an independent observation. An alternative way to study the disposition effect is to consider that

¹³ For a discussion on the limits of these measures, see for example Feng and Seasholes (2005)

¹⁴ Contrary to our simple illustration, in order to control for independence at an account level, the sale of a stock is counted only if no sale has been previously counted for that stock in any account within a week before or after the sale date (Odean, 1998).

¹⁵ Note that if many operations are recorded on the same day, a unique disposition effect is computed.

realized/paper gains and losses are independent, not at the transaction level, but at the account or investor level. In Table 4, we provide the average values of *PGR*, *PLR* and *DE* over 1 998 924 independent observations¹⁶.

Table 4 about here

On the entire sample, the null hypothesis ($PGR \leq PLR$) is rejected with a high degree of statistical significance. Investors are prone to the disposition effect over our sample period. Note that the results differ across years. For example, in 1999 the aggregate disposition effect is the highest (0.1078) whereas 2006 exhibits the lowest DE value (0.013).

However, looking at the evolution of the average DE and of the ratio PGR/PLR, we cannot highlight any distinct monotonic trend over time. For example, PGR/PLR values gradually increase from 2000 to 2002, peaking in 2003 and falling off as from 2004. The ratio PGR to PLR is the rate at which the individual investors prefer to sell winning stocks rather than losing ones. On the average, a stock that is up in value is more than 60% (1.68) more likely to be sold than a stock that is down. These results are quite in line with those generally obtained in the literature: Odean (1998) and Weber and Welfens (2006) compute a ratio of 1.5 while Brown *et al.* (2006) and Chen *et al.* (2007) get 1.6.

For a better understanding of the behavior of the investors, Table 5 (column 1) gives the average returns since the day of purchase for realized and paper gains and losses for the entire sample. Returns on paper gains are fourfold greater than those on realized gains. The same type of conclusion is obtained for losses (last two rows of Table 5). As noted by Odean (1998), these results seem to confirm that investors are more likely to realize smaller, rather than larger, gains and losses

Table 5 about here

¹⁶ We also use this methodology over our dataset (see figure 4 in the appendix) and find that approximately 20 % of the investors do not exhibit any DE or exhibit the opposite behaviour ($DE < 0$). This result confirms the ones obtained by Dhar and Zhu (2006) on US individual investors.

We also test whether the disposition effect observed in our sample can be explained by the desire of individuals to rebalance their portfolios (Table 6, column 1) or to restore diversification (Table 6, column 2). For the first test, we eliminate any sale for which the entire position in a stock has not been cleared (53 502 investors sold their entire position in the database). To eliminate any transaction resulting from a desire to restore diversification, we also remove sales for which there has been a new purchase on the sale date or during the 3 following weeks (21 days). 48 523 traders are concerned. Our results confirm previous results by demonstrating that traders still prefer to sell winners. The magnitude of the disposition effect is not reduced on this restricted sample.

Table 6 about here

In order to investigate the influence of traders' sophistication on the disposition effect, we build different groups of traders and check whether they exhibit any disposition effect. Three proxies for investors skills are retained; the geographical diversification of trades (presence of trades outside France), the use of the French SRD ("Système à Règlement Différé") and the investment in warrants. Briefly speaking, although individual investors are not usually supposed to be sophisticated ones, we assume that among them, those who internationally diversify portfolios (or are subject to a less important home bias), trade with SRD or trade warrants are at least more familiar with financial markets.

According to the ISIN of stocks, we divide investors in two categories: 40 430 among them only invest in French stocks, and we call them "local traders", the others are "international traders". Results in Table 7 panel I indicate that both groups are prone to the bias. More precisely, the disposition effect for "local traders" is 0.093 which is twofold the value of the disposition effect of "international traders".

Table 7 about here

Note that *PGR* and *PLR* measures are dependent on the portfolio size; we could obtain a lower disposition bias for an individual trading frequently but realizing the same number of winners/ losers. We should however point out that Dhar and Zhu (2006) compare measures of DE over sub groups. They justify such comparisons by the relative homogeneity of portfolios size among groups. Therefore, we computed the number of stocks held by individuals in each of our sub-groups: local traders have on average 14 securities while international investors hold 33 stocks. Given the difference in portfolio sizes, we do not compare our measures of DE¹⁷. The same argument applies to the other two proxies for sophistication (warrants and SRD use) although these proxies are not directly linked to trading behavior¹⁸ during the sample period as they rely on the presence of specific trades in each investor account. A “SRD” investor always chooses to use the leverage and short selling facility; there are 1 095 such investors. A “warrant” investor trades warrants at least once during the sample period; there are 11 460 such investors.

Results in Table 7, panels II (SRD) and III (warrants), indicate that the 4 groups are prone to the bias and that the disposition effect appears to be slightly lower for sophisticated traders (DE is 0.045 for SRD investors and 0.043 for warrant investors against 0.051 and 0.055 for the respective non sophisticated investors). Though more investigation is clearly needed, sophistication seems to attenuate the DE which order of magnitude is 0.04 for all sophisticated investors (0.048 for international traders)¹⁹.

IV – 2 Disposition effect and taxes

In this subsection, we first analyze the existence of end-of-the-year effect on the disposition effect (tax impact). Secondly, with respect to French specificities, we also investigate whether account types and tax regime shifts influence investment behavior.

¹⁷ For a discussion of DE determinants, see for example Feng and Seasholes (2005).

¹⁸ Note that “warrants” trades are excluded from our dataset.

¹⁹ We also use the trading activity (based on the number of annual transactions) as another proxy and find DE=0.04 for frequent investors. We do not report these results because we think trading activity constitutes a proxy for experience that does not always hint to sophistication.

In order to investigate whether individual investors pay attention to tax considerations at the end of the fiscal year, we also compute *PGR*, *PLR* and *DE* over the two intra year periods, January-November and December. Drawing on the work of Constantinides (1984), we expect investors to gradually increase their tax-loss selling from January to December. Table 8 provides the results.

Table 8 about here

We test the differences in proportions over the two sub-periods. Formally, for two independent samples (1) and (2), we test the following hypothesis:

H_0 : Proportion of Gains Realized in (2) = Proportion of Gains Realized in (1)

and

H'_0 : Proportion of Losses Realized in (2) = Proportion of Losses Realized in (1)

The following statistic (normally distributed) is applied to test H_0 where:

$$T_{H_0} = \frac{PGR_2 - PGR_1}{\sqrt{\hat{\pi}(1 - \hat{\pi}) \left(\frac{1}{(N_{RG_2} + N_{PG_2})} + \frac{1}{(N_{RG_1} + N_{PG_1})} \right)}}$$

$$\text{with } \hat{\pi} = \frac{(N_{RG_2} + N_{PG_2})PGR_2 + (N_{RG_1} + N_{PG_1})PGR_1}{(N_{RG_2} + N_{PG_2}) + (N_{RG_1} + N_{PG_1})}$$

N_{RG_j} and N_{PG_j} denote the number of realized gains and of potential gains (paper gains) in sample j .

In previous studies, *DE* is generally negative and the *PGR/PLR* ratio is lower than 1 in the last month of the fiscal year (December in US market for Odean (1998) and June for Brown *et al.* (2006) in Australia, for example).

In table 8, the disposition effect seems to be lower in December when compared with the average value of January-November but it is still positive. Tests of differences in proportions

indicate that the following results are significant: $PGR_{Jan-Nov} > PGR_{Dec}$ and $PLR_{Dec} > PLR_{Jan-Nov}$. These tests show that the lower DE in December is due to an average lower PGR and a higher PLR in December. This result differs from Odean's conclusion of a lower DE in December which was due to both significantly higher PLR and PGR in December.

Moreover, looking in table 8 at PGR/PLR indicates that on average, traders realize fewer gains and more losses in December: the ratio of PGR over PLR being 1.57 in December against 1.68 for the entire year. However, the fiscal impact in France appears to be moderate relative to the one observed in other countries as PGR/PLR remains higher than one in December.

The results in Table 5 (column 2 and 3) also help to confirm the presence of a moderate fiscal impact at the end of the year. Returns on realized paper losses are $-0,079$ in December against $-0,068$ for the entire year. These results are clearly different from Odean ones who obtain a greater difference between these two values ($-0,366$ in December against $-0,228$).

Finally, Figure 2 plots the average ratio of PGR/PLR on a monthly basis. We notice that contrary to Constantinides' (1984) arguments, investors do not gradually decrease the rate at which they sell winning stocks compared to losing ones during the year.

Figure 2 about here

In the French case, the fiscal impact on the selling behavior of investors could be tested in an original way due to the tax regime of some accounts (PEA accounts). Actually, capital gains are tax-free for all trades occurring 5 years after the opening date of the account. To be more precise, it is important to understand that fiscal exoneration occurs even if stocks were not kept for more than 5 consecutive years. The only legal restriction imposed before 5 years is that investors can't withdraw cash resulting from sales. For example, a capital gain on a round-trip trade made five years after the inception date of the account is tax-free.

Therefore, as investors may choose to trade on both accounts, we expect to measure the impact of tax on selling behavior by highlighting different behaviors on PEA accounts and traditional accounts. To serve our purpose, we focus our analysis on investors trading both on

PEA and traditional accounts²⁰. In this context, for any holder of a PEA account, 5 years represents a focal point (beginning of the tax-free period). If investors are sensitive to taxes, we expect buy and sell decisions to be affected by the tax shift on the PEA account after 5 years. To control our results, we study the same behavioral patterns for the same investors on traditional accounts.

We identify traders holding more than five years old PEA and traditional accounts (2 116 investors that we call “GROUP I”) and classify trades made on these accounts according to their execution date. In other words, we distinguish trades that were realized before and those realized after the accounts reached the focal point of five years. This ensures a good comparative basis for any analysis of possible different behaviors.

Table 9 gives the results obtained for the 2 116 investors (Group I) at an aggregate level. Global results indicate that the disposition effect is clearly positive and significant before and after five years on both accounts. Accurately, on traditional accounts, the DE before five years is 0.076 (column 1) and 0.034 for trades made after five years (column 3). For PEA accounts values are 0.084 (column 2) and 0.032 (column 4). At an aggregate level we observe that DE decreases between the two sub-periods whatever the account type²¹. Figures 3 confirms this result and gives a more precise illustration of the evolution of the aggregate DE with respect to experience (years of trading) for the 2110 investors. For example, at the end of the second year of trading DE is 0,056 and at the end of the seventh year of trading the value is about 0,02. This curve could be seen as an “experience curve” and the decreasing trend could be linked to the role played by the number of years of trading ; the impact of this variable was demonstrated in previous studies (Dhar and Zhu, 2006, Shu *et al.* 2005, Brown *et al.*, 2006 for example)²².

Table 9 about here

²⁰ On the entire sample, there are 35 598 such traders. Note that there are 46 094 holders of only traditional accounts and 10 911 holders of only PEA in the database.

²¹ Note that results for the 1665 investors trading only on PEA accounts and keeping this account for more than five years (Group II) and for the 5114 investors trading only on traditional accounts and keeping this account for more than five years (Group III) confirm this point (see table 12 in the appendix).

²² Note that the decrease of DE is essentially imparted to the decrease of PGR, investors seems to correct this bias in an asymmetrical way.

To investigate more accurately the hypothetical tax impact on selling behaviors, we compute the disposition effect at an individual level for the 2 116 investors belonging to Group I before and after the 5th birthday of both accounts.

The results for these investors are given in Table 10. Accurately, on traditional accounts, the DE before five years is 0.159 and 0.1 for trades made after five years. For PEA accounts values are 0.179 and 0.101. This table confirms the decrease of the individual average disposition effect after 5 years on the two account types and again highlights the results obtained at the aggregate level.

However, to control for any global compensation between investors, we conduct a Wilcoxon signed rank test of individual DE differences.

This test uses both the information on the direction and the relative magnitude of the differences within pairs of an identical trader average DE. For two distributions X and Y, the null hypothesis of the test is the following:

H₀: X and Y are samples from populations with same continuous distributions.

Table 10 about here

Figure 3 about here

Table 11 gives the results of the tests for the differences in distributions between types of accounts and detention duration. We denote (A) [resp. B] the distribution of the individual DE for trades over PEA before 5 years [resp. after 5 years] and (C) [resp. D] is the distribution of DE for trades on traditional account before 5 years [resp. after 5 years]. V is the number of ranks of positive differences. Note that as $N=2116$ is a large sample size, the number of the ranks of positive differences, V, is approximately normal.

The two first columns (A/B and C/D) show that individual distributions before and after are significantly different given account types. The behavior of investors seems to be clearly different as experience increases; this confirms the importance of learning already highlighted

at an aggregate level (see figure 3). The test on B/D distributions allows us to reject the tax argument for the PEA account. Actually, in the period of different taxation between both accounts, no difference of trading behavior in any direction could be detected at an individual level.

Table 11 about here

V Conclusion

We provide first and original results on the behavior of investors in the French context. On a large and proprietary anonymous database provided by Cortal Consors, a French broker, we find strong evidence that the disposition effect is observed for different categories of investors and for all time periods. Moreover this mistaken behavior does not seem to be motivated by a desire to rebalance portfolios.

As we expect some particular traders to be more sophisticated than others, based on original proxies (international diversification, SRD use, for example) we demonstrate that sophistication does not eliminate the existence of a disposition bias.

In this paper we conduct two tests of the impact of taxes on the selling behavior. First, at an aggregate level we find that investors are less prone to the disposition effect in December than during the rest of the year (due to a higher PGR and a lower PLR). Moreover, investors seem to realize losses of slightly stronger magnitude in December. However, unlike previous studies, DE is still positive (and PGR/PLR is higher than 1) in the last month of the fiscal year (Odean (1998), Brown *et al.* (2006), for example). Secondly, an analysis of a French specificity, i.e. the existence of tax free accounts (PEA more than 5 years old) allows us to demonstrate that accounts tax regimes have no impact on selling behavior.

Finally, this work could be extended at least in order to highlight characteristics of individual investors explaining the level of the disposition effect and its dynamics.

References

Barber, B. M., Y. T. Lee, Y. J. Liu and T. Odean. 2007. Is the Aggregate Investor Reluctant to Realize Losses? Evidence from Taiwan. *European Financial Management*, **13**(3), 423–447.

Barberis, N. and W. Xiong. 2006. What Drives the Disposition Effect? An Analysis of a Longstanding Preference Based Explanation. Forthcoming *Journal of Finance*.

Barberis, N. and W. Xiong. 2008. Realization Utility. Working Paper, Yale University.

Brockner, J. 1992. The Escalation of Commitment to a Failing Course of Action: Toward Theoretical Progress. *The Academy of Management Review*, **17**, 39-61.

Brown, P., N. Chappel, R. da Silva Rosa, and T. Walter. 2006. The reach of the disposition effect: Large sample evidence across investor classes, *International Review of Finance*, **6**(1-2), 43-78.

Chen, G.M, K. A. Kim, J. R. Nofsinger and O. M. Rui . 2004. Behavior and performance of emerging market investors: Evidence from China. Working paper.

Chui, P. 2001. An Experimental Study of the Disposition Effect: Evidence From Macau. *The Journal of Psychology and Financial Markets*, **2**(4), 216-222.

Constantinides, G. 1984. Optimal Stock Trading with Personal Taxes: Implications for Prices and the Abnormal January Returns. *Journal of Financial Economics*, **13**, 65-89.

Coval, J. D. and T. Shumway 2005. Do Behavioral Biases Affect Prices? *Journal of Finance*, **60**(1), 1-34.

Dhar, R. and N. Zhu. 2006. Up Close and Personal: Investor Sophistication and the Disposition Effect. *Management Science*, **52**(5) 726-740.

Feng, L. and M. Seasholes 2005. Do Investor Sophistication and Trading Experience Eliminate Behavioral Biases in Financial Markets ? *Review of Finance*, **9**, 305-351.

- Ferris, S., R. Haugen and A. Makhija. 1988. Predicting Contemporary Volume with Historic Volume at Differential Price Levels: Evidence Supporting the Disposition Effect. *Journal of Finance*, **43**(3), 677-697.
- Frino, A., D. Johnstone and H. Zheng. 2005. The Propensity for Local Traders in Futures Markets to Ride Losses: Evidence of Irrational or Rational Behavior? *Journal of Banking and Finance*, **28**, 353-372.
- Genesove, D. and C. Mayer. 2001. Loss Aversion and Seller Behavior: Evidence from the Housing Market. *The Quarterly Journal of Economic*, **116**(4), 1233-1260.
- Grinblatt, M., et M. Keloharju. 2001. What Makes Investors Trade? *Journal of Finance*, **56**, 589-616.
- Heath, C., S. Huddart and M. Lang. 1999. Psychological Factors and Stock Option Exercise. *The Quarterly Journal of Economics*, **114** (2), 601-627.
- Hens, T. and M. Vlcek. 2005. Does Prospect Theory Explain the Disposition Effect? NHH Dept. of Finance & Management Science, Discussion Paper No. 18.
- Huberman, G. 2001. Familiarity breeds investment. *Review of Financial Studies*, **14**, 659-680.
- Kahneman, D. and A. Tversky. 1979. Prospect Theory: an Analysis of Decision under Risk. *Econometrica*, **47**, 263-291.
- Kyle, S.A., H. Ou-Yang and W. Xiong. 2006. Prospect Theory and Liquidation Decisions. *Journal of Economic Theory*, **129**, 273-288.
- Lakonishok, J. and S. Smidt. 1986. Volume for Winners and Losers: Taxation and Other Motives for Stock Trading. *Journal of Finance*, **41**(4), 951-974.
- Locke, P. and S. Mann. 2003. Professional Trader Discipline and Trade Disposition. Working Paper, George Washington University.
- Muermann, A. And J.M. Volkman. 2006. Regret, Pride and the Disposition Effect. Working Paper, Wharton School.
- O'Curry Fogel, S. and T. Berry. 2006. The disposition effect and individual investor decisions: the roles of regret and counterfactual alternatives. *Journal of Behavioral Finance*, **7**(2), 107-116.
- Odean, T. 1998. Are Investors Reluctant to Realize Their Losses? *Journal of Finance*, **53**(5), 1775-1798.

- Rubaltelli, E., S. Rubichi, L. Savadori, M. Tedeshi and R. Ferretti. 2005. Numerical Information Format and Investment Decisions: Implications for the Disposition Effect and the Status Quo Bias. *The Journal of Behavioral Finance*, **6**(1), 19-26.
- Shapira, Z. and I. Venezia. 2001. Patterns of Behavior of Professionally Managed and Independent Investors. *Journal of Banking and Finance*, **25**(8), 1573-1587.
- Shefrin, H. and M. Statman. 1985. The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence, *Journal of Finance*, **40**, 777-790.
- Shu, P.G., Y.H. Yeh, S.B. Chiu and H.C. Chen. 2005. Are Taiwanese individual investors reluctant to realize their losses? *Pacific Basin Finance Journal*, **25**(2), 201-223.
- Staw, B.M. 1979. Knee-deep in the Big Muddy: a Study of Escalating Commitment to a Chosen Course of Action. *Organizational Behavior and Human Performance*, **16**, 27-44.
- Thaler, R. H. and E. J. Johnson. 1990. Gambling with the House Money and Trying to Break Even; the Effects of Prior Outcomes on Risky Choice. *Management Science*, **36**(6), 643-660.
- Tversky, A. and D. Kahneman. 1992. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, **5**, 232– 297.
- Weber, E. and Hsee, C. (1998). Cross-cultural differences in risk perception but cross-cultural similarities in attitudes towards risk. *Management Science*, **44**, 1205-1217.
- Weber, M. and C. F. Camerer. 1998. The Disposition Effect in Securities Trading: An Experimental Analysis. *Journal of Economic Behaviour and Organization*, **33**(2), 167-184.
- Weber, M. and H. Zuchel. 2005. How Do Prior Outcomes Affect Risk Attitude? Comparing Escalation of Commitment and the House Money Effect. *Decision Analysis*, **2**(1), 30-43.
- Weber, M. and F. Welfens. 2006. An individual Level Analysis of the disposition Effect: Empirical and Experimental Evidence. Working Paper. University of Mannheim.
- Zuchel, H. 2001. Why Drives the Disposition Effect? Working Paper, University of Mannheim.

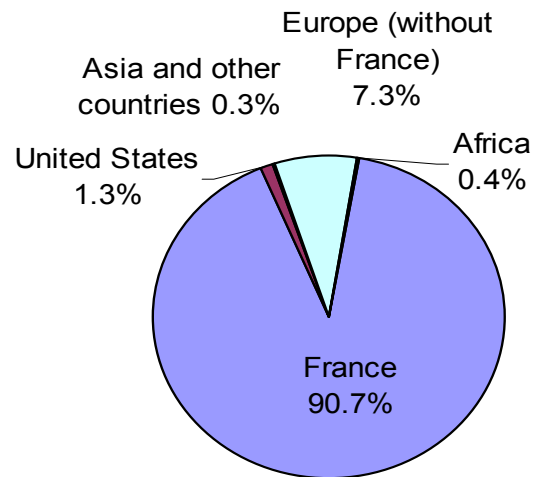


Figure 1 individual investors trades across the world

Table 1 Descriptive statistics on investors

This table contains results based on 8 438 885 trades (4 426 894 purchases and 4 011 991 sales) for 90 079 investors over 1999-2006. "Age" (in years) is computed on the 01/01/1999, "Activity over 1999-2006" is the number of investors active accounts : active accounts are those with at least one transaction over 2 years (consecutive or not). "trade amount / investor" [resp. Total Nb of trades/investor] is the total euro amount [Nb of trades] traded by investors over 1999-2006.

Variables	Mean	Std Dev.	25%	50%	75%	99%
Age	41.73	14.8	30	39	52	78
Assets / trade	460.24	4486.10	23	60	200	7000
Activity over 1999-2006	4.28	2.062	3	4	6	8
Trade amount/investor (€)						
Buy	3696.90	9373.90	1168.18	1961.28	4450.41	24299.19
Sell	4011.24	10387.02	1203.71	2188.86	4994.84	27327.46
Total Nb of trades/investor	93.68	354.45	6	22	74	1099

Table 2 Trades realized by investors

			Inv. I		Inv. II	
Dates	Asset	Price	Type	Nber	Type	Nber
10/1	X	25	B	40	B	50
	Y	10	B	50	B	65
	Z	30	B	20	NO	/
10/10	X	40	S	10	B	25
	Y	15	NO	/	B	15
	Z	20	NO	/	NO	/
10/31	X	45	S	10	S	40
	Y	8	S	5	S	10
	Z	10	NO	10	NO	/

Table 3 Key values for the two investors

	<i>RG</i>	<i>RL</i>	<i>PG</i>	<i>PL</i>
INV I				
10/10	1	0	1	1
10/31	1	1	0	1
Total I	2	1	1	2
INV II				
10/31	1	1	0	1
Total II	1	1	0	1
TOTAL (I +II)	3	2	1	3

Table 4 The disposition effect

This table contains results based on 4 011 991 sales over 1999-2006; 1 998 924 disposition effects are computed for 57 153 investors. N_{RG} , N_{PG} , N_{RL} , N_{PL} denote the number of realized gains, the number of potential gains (paper gains), the number of realized losses and the number of potential losses (paper losses). PGR (resp. PLR) denotes the proportion of realized gains (resp. the proportion of realized losses ratios). DE “disposition effect” is defined as $PGR - PLR$

	Entire Sample	1999	2000	2001	2002	2003	2004	2005	2006
N_{RG}	2044740	270763	484745	192737	133635	199390	199570	257344	306556
N_{PG}	14408013	1046046	2705794	1215168	798559	1167386	1687904	2669883	3117273
N_{RL}	1361264	96310	309986	211637	151656	134930	147961	137532	171251
N_{PL}	17076433	880578	3271772	2644645	2174947	2157664	1974155	1893423	2079250
PGR	0,124	0,206	0,152	0,137	0,143	0,146	0,10	0,088	0,089
PLR	0,073	0,098	0,086	0,074	0,065	0,059	0,069	0,067	0,076
PGR / PLR	1,68	2,10	1,77	1,85	2,2	2,47	1,45	1,31	1,17
DE	0,050	0,1078	0,065	0,063	0,078	0,087	0,036	0,020	0,013
Z-stat	496,51	230,82	261,63	191,15	196,71	256,27	126,79	83,52	57,29

Table 5 Average returns

	Entire sample	Jan-Nov	Dec
Return on realized gains	0.1116449	0.1116082	0.1120379
Return on paper gains	0.4019417	0.4066277	0.3517965
Return on realized losses	-0.0681329	-0.0670614	-0.0795994
Return on paper losses	-0.2421513	-0.2424635	-0.2388105

Table 6 Portfolio rebalancing

This table contains results based on 4 011 991 sales over 1999-2006; 1 998 924 disposition effects are computed for 57 153 investors. First column contains results when transactions associated to a sold of entire position are keep. Second column contains result when sales for which there has been a new purchase on the sale date or during the 3 following weeks are removed. *PGR* (resp. *PLR*) denotes the proportion of realized gains (resp. the proportion of realized losses ratios). DE “disposition effect” is defined as $PGR - PLR$.

	Entire Position sold	No purchase 3 weeks after sale
PGR	0.143	0.132
PLR	0.083	0.080
PGR / PLR	1.72	1.65
DE	0.060	0.052
Z-stat	400.43	177.79

Table 7 DE for groups

This table contains results based on 4 011 991 sales over 1999-2006; 1 998 924 disposition effects are computed for 57 153 investors. “Local” column contains results for 41 272 investors who only invest in French stocks. “SRD” column contains results for 1 095 investors who only use the SRD French system. “Warrant” column contains results for 11 460 investors who invest in warrants. N_{RG} , N_{PG} , N_{RL} , N_{PL} denote the number of realized gains, the number of potential gains (paper gains), the number of realized losses and the number of potential losses (paper losses). PGR (resp. PLR) denotes the proportion of realized gains (resp. the proportion of realized losses ratios). DE “disposition effect” is defined as $PGR - PLR$.

	Panel I		Panel II		Panel III	
	Local	International	SRD	Others	Warrant	Others
N_{RG}	174772	1869968	16980	2027760	746802	1297938
N_{PG}	676740	13731273	53819	14354194	5832682	8575331
N_{RL}	103683	1257581	10957	1350307	525914	835350
N_{PL}	824921	16251512	45606	17030827	6949025	10127408
PGR	0,205	0,119	0,239	0,124	0,114	0,131
PLR	0,111	0,071	0,194	0,073	0,070	0,076
PGR / PLR	1,85	1,67	1,23	1,69	1,63	1,72
DE	0,093	0,048	0,045	0,051	0,043	0,055
Z-stat	171,34	467,24	19,96	495,23	278,266	412,057

Table 8 DE over intra year periods

This table contains results based on 4 011 991 sales over 1999-2006; 1 998 924 disposition effects are computed for 57 153 investors. The data are partitioned into 3 different year periods; entire year, [January-November], December. *PGR* (resp. *PLR*) denotes the proportion of realized gains (resp. the proportion of realized losses ratios). DE "disposition effect" is defined as $PGR - PLR$.

	Entire Year	Jan-Nov	December
PGR	0.124	0.125	0.124
PLR	0.073	0.073	0.079
PGR/PLR	1.68	1.71	1.57
DE	0.050	0.051	0.044
Z-stat	496.51	482.32	134.91
T_{H_0} for PGR	-5.374	2.314	-0.011
T_{H_0} for PLR	2.966	-28.193	-27.030

Figure 2 Monthly level of PGR/PLR

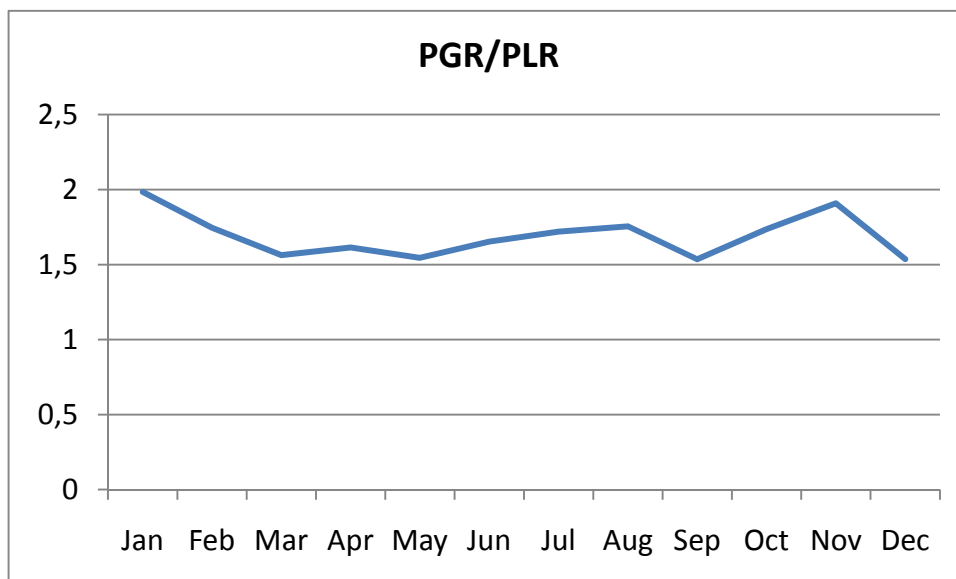


Table 9: DE before and after 5 years for group I (Aggregate DE)

This table contains results for investors trading simultaneously on PEA and traditional accounts and holding both accounts more than five years. Transactions are classified in two categories (realized before or after five years). N_{RG} , N_{PG} , N_{RL} , N_{PL} denote the number of realized gains, the number of potential gains (paper gains), the number of realized losses and the number of potential losses (paper losses). PGR (resp. PLR) denotes the proportion of realized gains (resp. the proportion of realized losses ratios). DE “disposition effect” is defined as $PGR - PLR$.

	CPT	PEA	CPT	PEA
	< 5 years	< 5 years	> 5 years	> 5 years
N_{RG}	79081	67015	91137	62455
N_{PG}	472676	419688	827693	758425
N_{RL}	47924	35845	59222	37382
N_{PL}	664230	637128	855285	816969
PGR	0.143	0.138	0.099	0.076
PLR	0.067	0.053	0.065	0.044
PGR/PLR	2.130	2.585	1.532	1.739
DE	0.076	0.084	0.034	0.032
Z-stat	136.410	149.510	85.156	88.115

Table 10: DE at an individual level before and after 5 years for group I

This table contains results for investors trading both on PEA and traditional accounts and holding both accounts more than five years. Transactions are classified in two categories (realized before or after five years) and DE is computed at an individual level. *PGR* (resp. *PLR*) denotes the proportion of realized gains (resp. the proportion of realized losses ratios). DE “disposition effect” is defined as $PGR - PLR$.

	Numbers	Mean	Standard deviation
PEA < 5 years (A)	2116	0.179	0.211
PEA > 5 years (B)	2116	0.101	0.192
CPT < 5 years (C)	2116	0.159	0.218
CPT > 5 years (D)	2116	0.100	0.220

Figure 3: Aggregate level of DE / year of trading

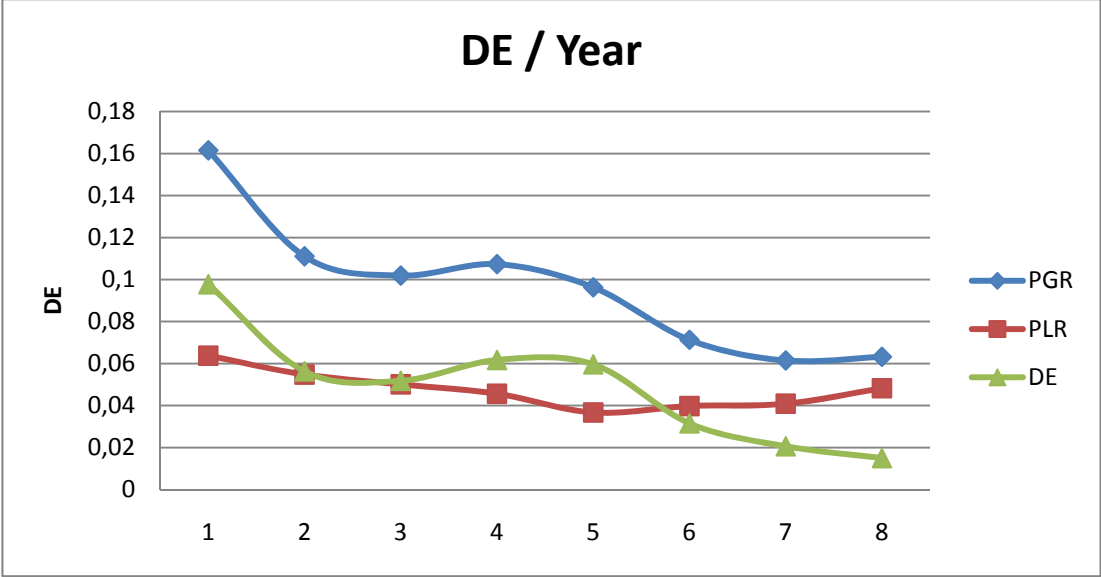


Figure 4: DE distribution at an individual level for investors holding PEA and traditional accounts before and after 5 years.

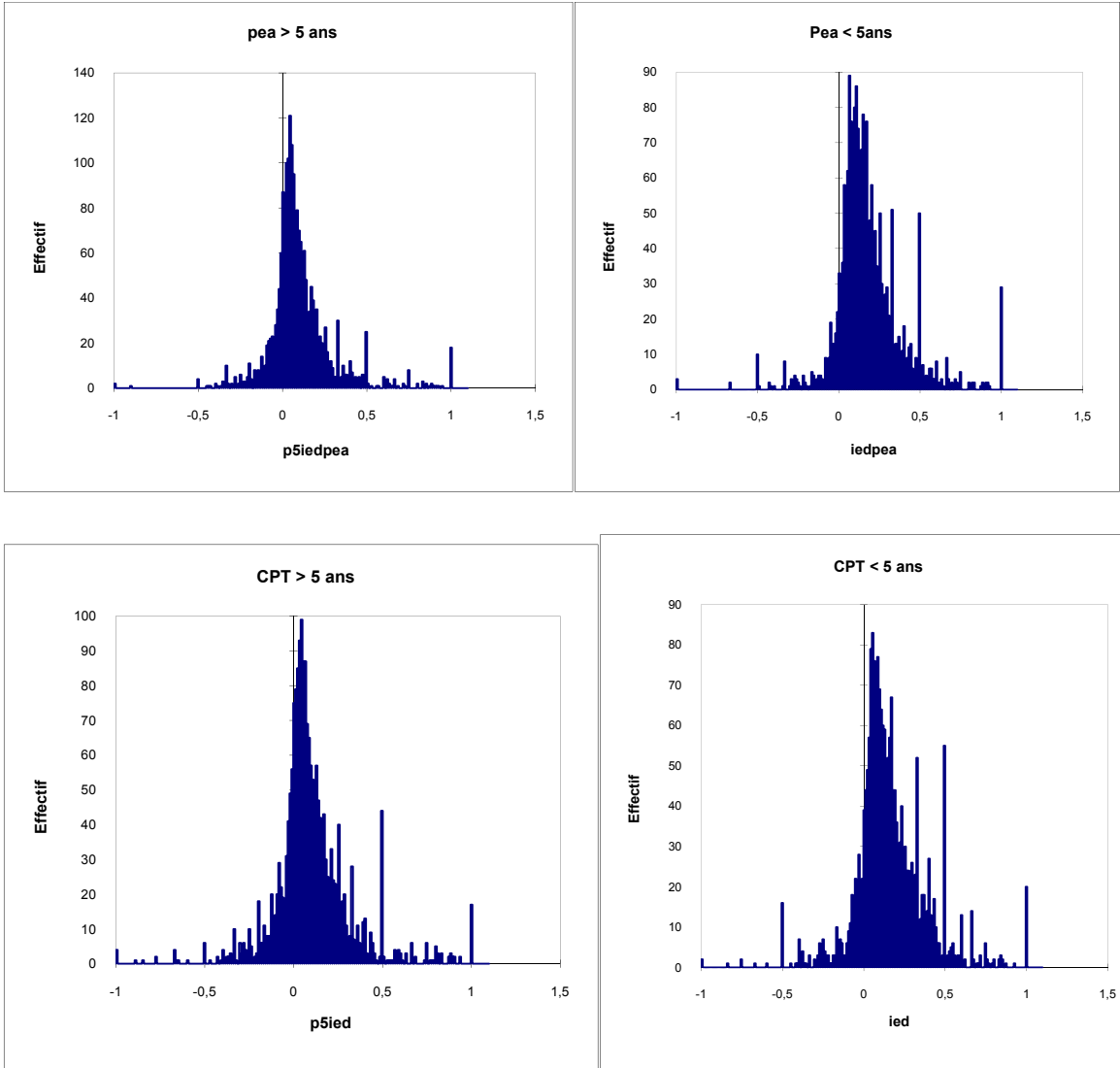


Table 11: Wilcoxon signed rank test for the differences in distributions A, B, C and D.

This table contains results for Wilcoxon signed rank test for investors trading on PEA and traditional accounts and holding both accounts more than five years. (A) [resp. B] denotes the distribution of individual level of DE for trades over PEA before 5 years [resp. after 5 years]. (C) [resp. D] denotes the distribution of DE for trades on traditional account before 5 years [resp. after 5 years]. V is the number of the ranks of positive differences.

	A/B	C/D	A/C	B/D
V	1680851	1465308	1228034	1117178
E(V)	1119870,5	1119882,5	1119870,5	1119891,5
Variance (V)	790084413,7	790084443	790084408,25	790084497,375
p-Value (Bilateral)	< 0,0001***	< 0,0001***	< 0,0001***	0,923
Alpha	0,05	0,05	0,05	0,923

Figure 4 Distribution of Disposition Effect for all Investors

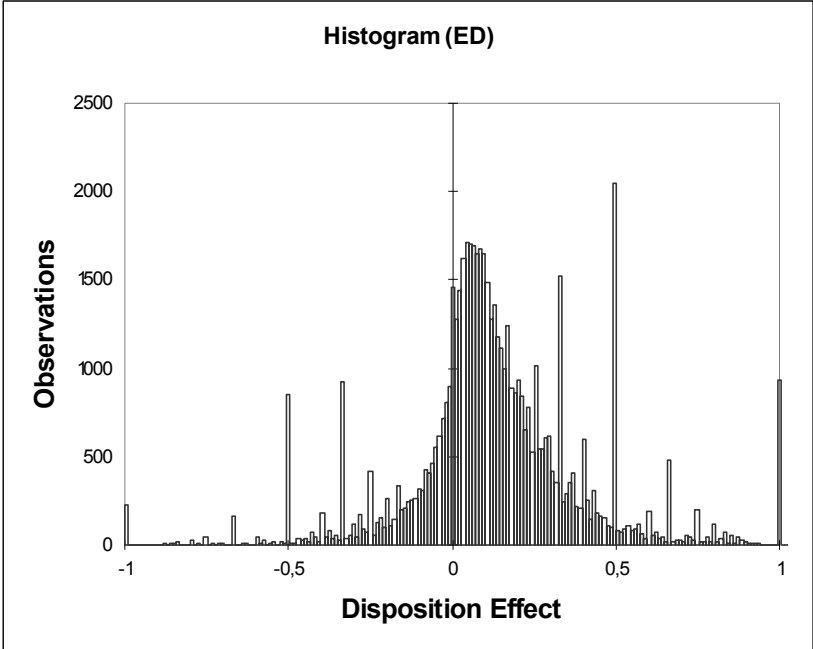


Table 11: DE for groups II and III

This table contains results for the 1665 investors trading only on PEA accounts and keeping this account for more than five years (Group II) and for the 5114 investors trading only on traditional accounts and keeping this account for more than five years (Group III). N_{RG} , N_{PG} , N_{RL} , N_{PL} denote the number of realized gains, the number of potential gains (paper gains), the number of realized losses and the number of potential losses (paper losses). PGR (resp. PLR) denotes the proportion of realized gains (resp. the proportion of realized losses ratios). DE "disposition effect" is defined as $PGR - PLR$.

	Group II		Group III	
	< 5 years	> 5 years	< 5 years	> 5 years
N_{RG}	37845	28981	463685	170254
N_{PG}	183956	245240	2263150	1473797
N_{RL}	21182	15773	337654	121196
N_{PL}	266288	238359	2825800	1489822
PGR	0,170	0,105	0,170	0,103
PLR	0,073	0,062	0,106	0,075
PGR/PLR	2,32	1,70	1,603	1,375
DE	0,096	0,043	0,063	0,028
Z-stat	103,60	57,58	221,22	89,73

Trading activity and Overconfidence: First Evidence from a large European Database

Boolell-Gunesh. S – Merli M²³.

March 2010

Abstract

We investigate the presence of overconfidence for 43 958 individual investors using a large brokerage account database between 1999 and 2006. We employ three methodologies to gauge overconfidence and our main results show that independently of the methodology considered, individual investors are subject to overconfidence and consequently trade too frequently. Securities investors are buying are systematically underperforming those they are selling on follow-up periods; investors are clearly not making profitable trades.

JEL Classification : G 10

Résumé

Nous étudions la présence de surconfiance pour 43 958 investisseurs individuels à partir d'une base de données de transactions individuelles sur la période 1999-2006. Trois méthodes visant à évaluer la rentabilité des stratégies d'investissement retenues par les investisseurs sont mises en œuvre et nos résultats montrent que quelle que soit la méthodologie considérée, les investisseurs individuels sont surconfiants et échangent trop fréquemment. Les stratégies adoptées par ceux-ci sont en moyenne sous-optimales, les titres achetés sous-performant systématiquement les titres vendus sur différents horizons d'investissement.

JEL Classification : G 10

²³ LaRGE / EM Strasbourg Business School, Strasbourg University, 61 avenue de la forêt noire, 67085 Strasbourg Cedex, France E-mail: shaneera@unistra.fr / merli@unistra.fr. We acknowledge the financial support of the European Savings Institute.

I. Introduction

Trading volume appears high on financial markets. For example, latest figures published by NYSE-Euronext in 2009 acknowledge that on average, 1.5 million securities have been traded each day in Europe. While this volume seems disproportionate to investors' rebalancing and hedging needs, proponents of behavioral finance suggest that one possible explanation to this phenomenon is overconfidence (see De Bondt and Thaler, 1995).

Overconfidence is a concept borrowed from psychology. In a broad manner, overconfidence can manifest itself in the following forms: miscalibration of probabilities, better than average effect, illusion of control and unrealistic optimism.

Miscalibration is the difference between the accuracy rate and probability assigned that an answer is correct. Overconfidence is a particular form of miscalibration in which the assigned probability that the answer is correct exceeds the true accuracy of answers. Overconfidence in terms of miscalibration is generally measured by asking people to construct confidence intervals for several uncertain quantities. The usual finding is that people's probability distributions are too tight and that the assigned probability is greater than the proportion of correct answers (Lichtenstein *et al.*, 1982). Concerning stock markets, it has been found that when people are asked to define confidence intervals for the return of an index or a stock, they usually underestimate volatility (DeBondt, 1998, Hilton, 2001, Glaser *et al.*, 2005)

The better than average effect hints to the fact that people believe that they are above average and that individuals have unrealistic positive views of themselves (Taylor and Brown, 1988, Cooper *et al.*, 1988). One famous example comes from Svenson (1981). While asking a group of US students to grade their driving skills, the author highlights that 50% of the participants rank themselves among the 30% of drivers with the highest driving safety.

Illusion of control relates to the observation that people tend to believe they can influence events which are in fact governed mainly by chance. An example of this illusion is given by Langer (1975), Miller and Ross (1975) or Presson and Benassi (1996) ; they show that people are ready to pay a higher price for a lottery ticket they can choose themselves, as if this could yield a more favorable result. Moreover, if these individuals expect certain

outcomes and these outcomes do occur, they are prone to assign them to their doing rather than to luck and reaffirm they can control events purely random events. Finally, overconfidence can lead to a form of unrealistic optimism. People believe that they have better chances than others to be confronted to positive events in their future life and particularly for their financial investments (Benartzi, Kahneman and Thaler, 1999). On the other hand, they think that others are more likely to face unfavorable events such as burglary or accidents.

Odean (1999), Barber and Odean (2001), Barber and Odean (2002) and Glaser and Weber (2007) show that overconfidence can lead to excessive trading on financial markets. Odean (1999) analyzes trading records of 10 000 individual investors in the US. The author considers that overconfidence relates to the miscalibration of the accuracy or precision of financial information. He formulates the following hypothesis. Over defined horizons, the securities these investors buy should outperform those they sell by at least enough to cover transaction costs. On the opposite, an overconfident trader who, for example, overestimates the precision of his information, may trade even if transaction costs are not taken care of. The author thus tests for overconfidence by determining whether, over appropriate horizons, securities investors buy outperform securities they sell by at least enough to cover transaction costs. The author shows that this is clearly not the case. For example, even before considering transaction costs, for the entire sample over a one year horizon, the average return to a purchased security is 3.3% lower than the average return to a security sold. Investors are not making profitable trades; they should trade less. Barber and Odean (2001) investigate overconfidence by using a “gender approach” ; they show that men are more overconfident than women. This behavior reduces net returns earned by men by 2.65% per year. Barber and Odean (2002) obtain similar results by analyzing the trading behavior of individual investors who switch to online trading. These investors earned exceptional returns in the period preceding their online debuts, beating the market by an average return of 2.4% on a yearly basis. The authors show that after going online, these investors trade more actively and less profitably; they underperform the market. Overconfidence can explain these observations: when people succeed, they often give themselves too much credit for their success. Failures, on the other hand, are blamed to others or to misfortune. It is likely that these investors thought that the excellent returns they earned before going online were due to their investment skills. Being “aware” of their talent, they increased their trading activity after going online and suffered from poor portfolio performance.

Finally, Glaser and Weber (2007) investigate the link between overconfidence and trading volume in an innovative way. The authors ask approximately 3000 online broker investors to answer an internet questionnaire designed to measure the following forms of overconfidence : miscalibration, the better than average effect, illusion of control and unrealistic optimism. Simultaneously, they compute several measures of trading volume of these individual investors (number of trades, turnover, etc). In a last step, they evaluate the correlation between the overconfidence scores they observe and the measures of trading volume. For instance, Glaser and Weber (2007) gauge “miscalibration” by asking general knowledge and stock market forecasts questions. Participants are asked to provide confidence intervals when answering questions. The “better than average effect” is measured by asking questions concerning skills and performance relative to others. An example of such a question is “*What % of customers of your discount brokerage house have better skills than you at identifying stocks with above average performance in the future?*” Results show that individual investors who think they are above average in terms of investment skills or past performance trade more. However, surprisingly, measures of miscalibration are unrelated to measures of trading volume. This conclusion is particularly striking because researchers incorporating overconfidence typically rely on the miscalibration literature to motivate this assumption.

However, even if the presence of overconfidence has been studied in a number of countries, no such research has yet been carried out in France. Furthermore, the only European empirical research dealing with the subject on individual data concerns 3 079 accounts (Glaser and Weber, 2007). Our paper fits this loophole by investigating the trading records of 43 958 individual investors at a French discount brokerage house between 1999 and 2006 and thus is the most comprehensive in the European context. The paper is organized as follows. Section II describes the data and introduces the methodology. Section III is devoted to the description of our main results and comments and section IV brings the concluding remarks.

II – Data and methodology

The main data is provided by a large French discount brokerage house. We obtain transaction records for all active²⁴ accounts over the period 1999-2006, that is 9 619 898 trades with 5 074 732 buy orders and 4 545 166 sell orders for 92 603 individual investors. The data is contained in three files: trades, investors and fees. The trades files combines the following data for each trade : ISIN code of the asset, type of asset (common stocks, certificates, warrants), buy-sell indicator, short sale indicator, quantity and amount in Euros, account type (traditional versus tax-free account or French Plan d'Epargne en Actions (PEA)) and the media used to place the order. The investors file gathers demographical information about the investors: date of birth, sex, place of living, among others. The fees file contains monthly fees paid by each investor and indicates whether they are trade fees or short sales fees.

In order to test for overconfidence, we extract a dataset containing only trades on European common stocks. This dataset includes 4 714 702 trades, with 2 482 190 buy orders and 2 232 512 sell orders made by 43 958 investors over 2 031 assets. For each stock we build a file containing historical daily prices over the period 1999-2006. In this respect, securities' ISIN codes are used to collect price data and information on splits and dividends through SIX-Telekurs -Fininfo²⁵.

The typical investor is a man (78.28 %), 47.9 years old on average. Table 1 describes some preliminary statistics. The average number of assets per trade is 495. Over the whole dataset period, investors realize 107 trades on average amounting to 3 945 € per trade (3 784 € for buy orders and 4 106 € for sell orders). Most of the trades relate to French stocks (94%), followed by trades of stocks from the Netherlands (4,1%) and from Luxembourg (0.7%). Over 1999-2006, the most commonly traded French, Dutch and Luxembourgish securities are

²⁴ Over the period 1999-2005, active accounts are those with at least one transaction over 2 years (consecutive or not). For the last year of the sample, accounts are active if they hold at least one transaction over the entire year.

²⁵ <http://www.six-telekurs.fr/>

respectively Alcatel-Lucent, Stmicroelectronics and Arcelor-Mital. Finally, on average investors are active 4.83 years over 8.

Table 1 about here

To test for overconfidence, we use the methodology proposed by Odean (1999). Over defined horizons, the stocks bought by investors should outperform those they sell by at least enough to pay the costs of trading. On the other hand, overconfident traders may hold mistaken beliefs about potential gains and buy and sell stocks even if these trades do not cover transaction costs. Investors may express overconfidence in two different manners. Following Odean (1999), either they overestimate the precision of private information or/and they have biased interpretation of publicly available information. If traders are informed but overestimate the accuracy of this information, they are not likely to face losses beyond the loss of transaction costs. On the opposite, if instead (or in addition) to being overconfident in the precision of information, investors overestimate their ability to interpret information; they may incur average losses beyond transaction costs.

In order to test for overconfidence, we decide to test if investors overestimate their ability to interpret information and determine whether the securities bought by the investors in the dataset outperform those they sell when trading costs are ignored. Note that the average trading costs paid by the investors, excluding the bid-ask spread is about 0.25%.

We look at return horizons of 3 months, 6 months and 1 year. To calculate the average returns to securities bought (sold) by the investors over the T ($T = 3$ months, 6 months and 1 year) trading period subsequent to the purchase, we index each purchase (sale) transaction with a subscript i , $i = 1, \dots, N$. Each transaction is composed of a security j_i and a date t_i . If the same stock is bought (sold) in different accounts on the same day, each purchase (sale) is counted as a distinct observation. The average return to the securities bought and sold over the T trading period is:

$$R_{P,T} = \frac{1}{N} \sum_{i=1}^N \prod_{\tau=1}^T (1 + R_{j_i, t_i + \tau}) - 1 \quad (1)$$

where $R_{j,t}$ is the daily return for security j on date t . We compare the average returns on securities bought and sold to the average amount of trading costs paid by investors.

We test whether over the same horizons, the average returns to securities bought are less than the average returns to securities sold, ignoring transaction costs. More formally, we test :

$$H_0: \text{Average returns to securities bought} \geq \text{Average returns to securities sold}$$

This assumption compares the average return to stocks bought and sold over subsequent periods. Returns are averaged over trading histories of investors and across investors. Many securities are bought or sold on more than one date and may even be bought or sold by the same investors on the same day. Suppose for example an investor sells a stock at a given date t and that 2 months later, another investor sells the same stock. Returns on this stock over a 3 month period are not independent because the periods overlap for 2 months. Thus, any statistical test which requires independence of observations cannot be employed. Statistical significance is thus estimated by conducting a Wilcoxon signed rank test of differences. We look at each investor in the dataset separately and determine for each of the chosen horizons (T=3 months, 6 months, 1 year) the average returns earned by the investor on the stocks bought and sold. We thereby construct distributions of average returns on securities bought and sold. The Wilcoxon test allows to compare the different distributions. Note that this test uses both the information on the direction and the relative magnitude of the differences of returns earned by the same investor on the stocks he purchases and stocks he sells. The null hypothesis is the following:

$$H'_0: X \text{ and } Y \text{ are samples from populations with the same medians and the same continuous distributions.}$$

This first methodology allows the computation of equally-weighted average returns. In order to complete our results and compare the average returns on portfolios bought and sold, we additionally measure the weighted average returns on securities bought and sold. More precisely, we basically follow the same steps as the ones involved in the methodology described before except that at the end, we report weighted average returns. These returns on portfolio bought (sold) are computed in the following manner. Looking at the total amount of euros spent by investors on securities bought (sold) over the period 1999-2006, we value the *weight* of each buy (sell) transaction. The weight of the buy transaction j is:

$$P_j = \frac{\text{Total amount of euros spent by investors on transaction } j}{\text{Total amount of euros spent by investors on buy orders}}$$

The weighted average returns to securities bought (sold) then refers to the sum of the returns earned on each transaction j multiplied by the appropriate weights P_j . As before, we test :

$$H''_0: \text{Average weighted returns to securities bought} \geq \text{Average weighted returns to securities sold}$$

We estimate statistical significance by conducting a Wilcoxon sign rank test of difference.

Finally, to check the robustness of our results, we also gauge the profitability of purchases and sales by using a calendar time method. We construct calendar-time portfolios consisting of all purchase (sale) events during a “portfolio formation period” (3 months and 6 months). To be more precise, each time a purchase (sale) occurs during the formation period, we assign this security to the calendar-time portfolio. The same security may have been bought or sold several times during the “portfolio formation period”. If this is the case, each purchase is counted as a separate observation. Each position is weighed equally. Finally, we calculate the “Buy” (“Sell”) portfolio return for the calendar month subsequent to the formation period. Rolling forward the formation period by one month, a time-series of calendar-time portfolio returns for month $t+1$ is obtained.

Based on the calendar-time approach, we calculate two measures of performance. The first one is simply the average monthly calendar time return on the “Buy” portfolio minus the “Sell” Portfolio. We test whether the average return on the “Buy” portfolio is less than the one to the “Sell” portfolio. The second performance measure is Jensen’s alpha (Jensen, 1969). We perform the following regression:

$$R_{B,t} - R_{S,t} = \alpha_p + \beta_p (R_{M,t} - R_{f,t}) + \varepsilon_{p,t} \quad (2)$$

where

$R_{B,t}$ is the monthly return on the calendar-time portfolio based on purchases

$R_{S,t}$ is the monthly return on the calendar-time portfolio based on sales

$R_{M,t}$ is the monthly return on a market index, the DJ EURO STOXX 50

$R_{f,t}$ is the monthly returns on BTAN

β_p is the market beta and $\varepsilon_{p,t}$ the regression error term.

III Results

(i) Average returns to securities bought and sold

Table 2 report the principal results. Panel A presents the average returns to securities bought and sold for the three follow-up periods we consider ($T= 3\text{ months}, 6\text{ months}, 1\text{ year}$). Panel B gives results for the Wilcoxon test.

The results in Panel A show that for all three horizons we study the average returns to securities bought are lower than to securities sold. For example, for the whole sample, over a 6 month period, the average return to a purchased security is 3.105 % lower than to a security sold. These results are particularly striking: stocks investors are buying are underperforming stocks they are selling even before taking into account transaction costs. In other words, investors are clearly not making profitable trades: they not only are paying transaction costs to trade but they are also making poor portfolio choices. Overconfidence in the ability to interpret information may be a good explanation for these findings.

Panel B gives indications about the statistical significance of our results. We denote (A) [resp. (B)] the distribution of average returns earned by each individual investor on purchased (resp. sold) stocks over a 3 month period subsequent to the purchase (sale). (C) [resp. (D)] is the distribution of average returns on bought (resp. sold) securities over a 6 month period. Finally, (E) [resp. (F)] refers to the distribution of average returns on purchased (resp. sold) stocks over a 1 year horizon. V is the number of ranks of positive differences. As $N = 43\ 958$ is a large sample size, the number of the ranks of positive differences, V , is approximately normal. Results show that the distributions of average returns to bought and sold securities are significantly different, whichever period considered ($T= 3\text{ months}, 6\text{ months}, 1\text{ year}$), H'_0 is rejected at all three horizons at a 1% level.

(ii) Weighted average returns to securities bought and sold

Table 3 presents the results obtained for weighted average returns. Panel A reports the average weighted returns to securities bought and sold for the three follow-up periods ($T= 3\text{ months}, 6\text{ months}, 1\text{ year}$). Panel B gives results for the Wilcoxon test.

Table 2 about here

Table 3 about here

The results bring further evidence of overconfidence in the trading behavior of our individual investors. Panel A show that for all three horizons we study ($T= 3\ months, 6\ months, 1\ year$), the average weighted returns to securities bought are again lower than to securities sold. For example, for the whole sample, over a 6 month period, the average return to a purchased security is 1.926% lower than to a security sold. Investors are making poor portfolio decisions.

Statistical significance is estimated in Panel B.. We denote (A) [resp. (B)] the distribution of average weighted returns earned by each individual investor on purchased (resp. sold) stocks over a 3 month period subsequent to the purchase (sale). (C) [resp. (D)] is the distribution of average weighted returns on bought (resp. sold) securities over a 6 month period. Finally, (E) [resp. (F)] refers to the distribution of average weighted returns on purchased (resp. sold) stocks over a 1 year horizon. V is the number of ranks of positive differences. As $N = 43\ 958$ is a large sample size, the number of the ranks of positive differences, V , is approximately normal. Results show that the distributions of average weighted returns to bought and sold securities are significantly different, whichever period considered ($T= 3\ months, 6\ months, 1\ year$).

(iii) Calendar time portfolios

Table 4 reports the results. Panel A presents the average monthly calendar time return on the “Buy” portfolio, on the “Sell” Portfolio as well as the difference between the returns. Panel B gives the estimates of Jensen’s alpha. t-values are given in parentheses.

The results confirm the ones obtained in the first part of this section. The values reported in Panel A show that for the two portfolio formation periods, the monthly returns to

the “Bu /Sell” portfolio are significantly negative. For instance, if we consider a portfolio formation period of 6 months, the “Sell” portfolio outperforms the “Buy” portfolio on average by 0.2%. Investors in the dataset are clearly making poor portfolio choices ; they are trading too much.

Results from Panel B confirm these conclusions. The Jensen’s α estimates are all significantly negative. The portfolios held by our investors are underperforming the market, once more highlighting poor portfolio selection

IV. Conclusion

In this paper, we investigate the presence of overconfidence for 43 958 individual investors using a large brokerage account database between 1999 and 2006. Based on three methodologies, our main results show that individual investors are subject to overconfidence and consequently trading too frequently. More precisely, securities or portfolios investors are buying are systematically underperforming those they are selling on follow-up periods. This study confirms that investors are clearly not making profitable trades.

References

Barber, B., Odean, T., 2001, Boys will be Boys: Gender, overconfidence, and Common Stock investment, *Quarterly Journal of Economics*, 116, 261-292.

Barber B., Odean T., 2002. Online investors: Do the slow die first?, *Review of Financial Studies* 15, 455-488.

Benartzi, S. Kahneman, D. and Thaler, R., 1999, Optimism and overconfidence in asset allocation decision, *news.morningstar.com*

Cooper, R., Woo, C., and Dunkelberg, W., 1988, Entrepreneur’s Perceived Chances of Success, *Journal of Business Venturing*, 3, 97-108.

De Bondt W., Thaler R., 1995. *Financial decision making in markets and firms: A Behavioral Perspective*, in Handbook in Operations Research and Management Science, ed. by Jarrow V., Maksimovic R., Ziemba W., Elsevier.

De Bondt W., 1998. A Portrait of the Individual Investor, *European Economic Review*, 42, 831-844.

Glaser, M., Langer, T., Weber, M., 2005, Overconfidence of Professional and Lay Men : Individual Differences within and between Tasks, Working Paper, University of Mannheim.

Glaser, M., Weber, M., 2007. Overconfidence and Trading Volume. *The Geneva Risk and Insurance Review*, 32(1), 1-36.

Hilton, D. J., 2001, The psychology of Financial Decision-Making: Applications to Trading Dealing and Investment Analysis, *Journal of Psychology and Financial Markets*, 2, 1, 37-53

Langer E., 1975. The illusion of control. *Journal of Personality and Social Psychology*, 32, 311-328.

Lichtenstein S., Fischhoff B., Philips L., 1982. Calibration of probabilities: the state of the art to 1980, in *Judgment Under Uncertainty :Heuristics and Biases*, ed. by Kahneman D., Slovic P., Tversky A., Cambridge University Press.

Miller, D. T., and Ross, M., 1975, Self-Serving Biases in the Attribution of Causality: Fact or Fiction?, *Psychologic Bulletin*, 82, 213-225

Odean T., 1999. Do Investors Trade Too much? *American Economic Review* 89, 1279-1298.

Presson P., Benassi V., 1996. Illusion of control: a meta-analytic review, *Journal of Social Behavior and Personality*, 11, 493-510.

Svenson O., 1981, Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica*, 47, 143-148.

Taylor S., Brown J., 1988. Illusion and well-being: A social psychological perspective on mental health, *Psychological Bulletin* 103, 193-210.

Tableau 1 Descriptive statistics

This table contains results based on 4 714 702 trades (2 482 190 buy orders and 2 232 512 sell orders) for 43 958 investors over 1999-2006. “Age” (in years) is computed on 01/01/2006, “Activity over 1999-2006” is valued for active accounts. “Trade amount / investor” [resp. Total Nb of trades/investor] is computed over the total amount of euros [Nb of trades] traded by investors over 1999-2006.

Variable	Mean	Std. Deviation	25%	50%	75%	99%
Age	47.90	13.67	37	46	57	83
Nb. Assets/trade	495.50	5284	24	60	200	7700
Trade amount/trade (€)						
Buy	3784.51	9850	1199.78	2013.01	4642.86	24299.19
Sell	4106.44	10757	1241.43	2276.09	5243.12	27290.40
Total Nb. Of trades /investor	107	392	21	43	99	994
Activity over 1999-2006	4.83	2.03	3	5	7	8

Table 2 Average returns and statistical significance

This table contains results based on 4 714 702 trades. Panel A reports average returns (%) to securities bought and sold for 3 follow-up periods: 3 months, 6 months and 1 year. "N" refers to the number of observations and "Difference" reports the difference between the average returns to securities bought and sold over each horizon. Panel B gives indications about statistical significance of the results reported in Panel A. (A) ([resp. (B)]) refers to the distribution of average returns earned by each individual investor on purchased (resp. sold) stocks over a 3 month period subsequent to the purchase (sale). (C) [resp. (D)] is the distribution of average returns on bought (resp. sold) securities over a 6 month period. (E) [resp. (F)] refers to the distribution of average returns on purchased (resp. sold) stocks over a 1 year horizon. V is the number of ranks of positive differences.

Panel A : Average returns following purchases and sales (%)				
	N	3 months later	6 months later	1 year later
Purchases	2 482 190	2.085	2.249	2.569
Sales	2 232 512	4.434	5.354	6.697
Difference		-2.349	-3.105	-4.128
Panel B : Wilcoxon Tests				
	A/B	C/D	E/F	
V	239 357 494	242 735 973	233 872 295	
E(V)	482 999 517.5	482 845 689	478 067 292.5	
Variance (V)	707 6667 430 651.25	7073 286 972 442.5	6968 547 889 126.12	
p-Value (Bilateral)	< 0.0001***	< 0.0001***	< 0.0001***	
Alpha	0.05	0.05	0.05	

Table 3 Weighted average returns and statistical significance

This table contains results based on 4 714 702 trades. Panel A reports average weighted returns (%) to securities bought and sold for 3 follow-up periods: 3 months, 6 months and 1 year. "N" refers to the number of observations and "Difference" reports the difference between the average weighted returns to securities bought and sold over each horizon. Panel B gives indications about statistical significance of the results reported in Panel A. (A)[resp. (B)] refers to the distribution of average weighted returns earned by each individual investor on purchased (resp. sold) stocks over a 3 month period subsequent to the purchase (sale). (C) [resp. (D)] is the distribution of average weighted returns on bought (resp. sold) securities over a 6 month period. (E) [resp. (F)] refers to the distribution of average weighted returns on purchased (resp. sold) stocks over a 1 year horizon. V is the number of ranks of positive differences.

Panel A : Average weighted returns following purchases and sales (%)				
	N	3 months later	6 months later	1 year later
Purchases	2 482 190	0.6106	1.628	2.177
Sales	2 232 512	2.171	3.554	4.638
Difference		-1.555	-1.926	-2.462
Panel B : Wilcoxon Tests				
	A/B	C/D	E/F	
V	466 750 742,5	452 521 632,5	525 668 467,5	
E(V)	482 999 517,5	482 845 688,5	478 067 292,5	
Variance (V)	7076 667 430 553,37	7073 286 972 351,62	6968 547 889 042,37	
p-Value (Bilateral)	< 0,0001***	< 0,0001***	< 0,0001***	
Alpha	0,05	0,05	0,05	

Table 4 Calendar time portfolios

This table contains results based on 4 714 702 trades. Calendar-time portfolios consisting of all purchase (sale) events during a “portfolio formation period” (3 months and 6 months) are constructed and two measures of performance are computed. Panel A reports the average calendar time return on the “Buy” and “Sell” portfolio. “Difference” is the difference between the average returns on the two portfolios. Panel B gives of Jensen’s alpha. t-values are given in parentheses.

Panel A : Returns on “Buy” and “Sell” Portfolios		
	Portfolio formation period	
	3 months	6 months
“Buy” portfolio	-0.010	-0.009
“Sell” portfolio	-0.007	-0.008
Difference	-0.003**	-0.002*
	(-2.17)	(-1.553)
Panel B: Jensen’s alpha		
	Portfolio formation period	
	3 months	6 months
Jensen’s α	-0.260**	-0.135*
	(-2.496)	(-1.68)
β	0.118***	0.092***
	(7.56)	(7.67)

***, **, * - Significant at the 1%, 5% and 10% levels respectively

Testing alternative theories of financial decision making: a survey study with lottery bonds

January 2010

Abstract

We present the results of a simple, easily replicable, survey study based on lottery bonds. It is aimed at testing whether agents make investment decisions according to expected utility, cumulative prospect theory or optimal expectations theory, when they face skewed distributions of returns. We show that more than 55% of the 245 participants obey optimal expectations theory. They choose a distribution of payoffs which is dominated for second-order stochastic dominance and which would not be chosen according to cumulative prospect theory, for a large range of parameter values.

Our results first cast doubt on the relevance of variance as a measure of risk; they show the importance of skewness in decision making and, more precisely, they emphasize the attractiveness of the best outcome, an essential feature of optimal expectations theory. The ranking of outcomes, used in cumulative prospect theory, seems insufficient to characterize the way people distort beliefs. As by-products of this study, we illustrate that agents use heuristics when they choose numbers at random and have, in general, a poor opinion about the rationality of others.

JEL classification: D03, D81

Keywords: Lottery bonds, optimal beliefs, probability distortion, risk aversion

1. Introduction

Standard economic theory assumes relatively simple rules to describe human behavior. Agents are supposed to manage any quantity of information they receive, according to Bayes' rule, and to take decisions without emotions or distorted beliefs. Their objective is to maximize the expectation of their utility function (henceforth EU model). There is now some evidence that agents make systematic "errors", especially in assessing probabilities. The probability of very good outcomes tends to be overvalued (Alpert and Raiffa 1982, Buehler *et al.* 1994; Weinstein 1980), introducing an optimistic bias. The consequence is a sub-optimal allocation of wealth. For example, households' portfolios are not well diversified. A part of their wealth is invested in mutual funds (and then well diversified) but another part, in general non negligible, is concentrated on a few stocks (Calvet *et al.* 2007, Goetzmann and Kumar 2008, Mitton and Vorkink 2007, Polkovnichenko 2005). Moreover, portfolios are often biased toward lottery-type stocks with positive skewness. Barberis and Huang (2008), Bali *et al.* (2009), Kumar (2009) and Mitton and Vorkink (2007) recently published papers focused on that problem. Barberis and Huang (2008) show that stocks with positively skewed returns can be overpriced on markets populated by investors obeying cumulative prospect theory (CPT in the following). It is especially the case if the return on skewed securities is independent of the returns on other securities and if the supply of skewed stocks is small relative to the global market supply. Kumar (2009) shows the existence of significant links between investment behavior and lottery play behavior. He shows that investors used to play (unfair) state-lotteries also prefer lottery-like stocks. This observation is reinforced during economic downturns. Bali *et al.* (2009) show that stocks exhibiting at least one very high return in the past month are overpriced. This effect is robust when controlling for idiosyncratic volatility. Mitton and Vorkink (2007) analyze the behavior of more than 60,000 retail investors and show that those who are not diversified select a few highly skewed stocks. It then seems that gambling and investment behaviors cannot be disentangled because of the preference for skewness and/or the attractiveness of the best outcome.

Brunnermeier and Parker (2005) and Brunnermeier *et al.* (2007) developed the theory of optimal expectations (OET in the following) to take into account this optimistic bias. In the second paper, Brunnermeier *et al.* (2007) consider a simple one-period, two-date model; they assume that agents behave optimally given their beliefs, and choose portfolios maximizing the expected present value of future utility flows. Roughly speaking, the felicity of agents is composed of *ex ante* and *ex post* utility. *Ex ante*, it is optimal to distort beliefs in an optimistic

way. However, this distortion comes at a cost, lying in a sub-optimal portfolio choice and a lower *ex post* expected utility. The authors call “optimal beliefs” the subjective assessment of probabilities which maximizes an average of *ex ante* and *ex post* utilities. In a complete market framework with a finite number of states of nature, they show that the optimal portfolio contains the risk-free asset, and the most skewed asset. Concerning optimal beliefs, they prove that the probability of only one state is overvalued, the probabilities of the other states being undervalued.

In cumulative prospect theory, distortion of beliefs is linked to payoffs, only through the ranking of gains and losses. Consequently, in a finite state space, the outcomes of two comonotonic prospects are weighted identically and independently of the values of the outcomes; only ranking matters. Concerning the distortions of beliefs, our results show that a large proportion of agents not only take into account the ranking, but also the values of the outcomes, especially the largest one.

In this paper, we present a survey study to test whether the attractiveness of the best outcome is really an important component of the decision making process or if agents behave according to the EU or CPT models. The test is based on a questionnaire asking participants to choose among different random outcomes of lottery bonds. These securities are well suited to address the question we are dealing with. First, they exist in many countries for more than two centuries, and are, even today, very popular (Green and Rydqwist 1997, Guillen and Tschoegl 2002, Lévy-Ullmann 1896, Millar and Gentry 1980, Pfiffelmann and Roger 2005, Ridge and Young 1998, Tufano 2008). Second, contrary to the distribution of stock returns which is unknown, the distribution of lottery bond returns is, in general, perfectly known. The possible outcomes are given objective probabilities. Third, most people who never invested in lottery bonds easily understand how payoffs are defined because they bet, at least occasionally, on state lotteries like the lotto game²⁶.

We consider two designs for the lottery bonds. They differ by the way the amount distributed through the lottery is defined. The first bond is designed to study the decision making process and to answer our main question concerning the choice of EU, CPT or OET. The individual amount received by winners through the lottery is known in advance and the remaining amount to be shared among all subscribers (including the winners) is random. The second bond is aimed at controlling for the “minimum” required level of rationality, that is, the compliance to the first-order stochastic dominance principle. In this case, the global amount

²⁶ Skewness-seeking investors are labeled “Lotto investors” by Mitton and Vorkink (2007).

paid through the lottery is known in advance (but the individual gain is random due to a pari-mutuel feature). The remaining amount shared by all subscribers is then not random. In fact, as we cannot define an incentive compatible payment scheme for the participants (without assuming that one theory is better than the others), the questions related to this second bond allow to “select” respondents that provide answers compatible with first-order stochastic dominance.

Our results show that more than 55% of participants behave like OET investors, exhibiting a preference for the random payoff with the highest possible outcome. It is important to notice that the possible random payoffs of our first bond have the same expected value and that the variance of returns is the highest for the payoff with the largest outcome. Moreover, our design is such that the payoff including the highest possible outcome is dominated by all other choices in the sense of second-order stochastic dominance. It then reinforces our result in favor of OET.

Our analysis also provides two by-products. The first one concerns the random choice of numbers. We illustrate that, at the aggregate level, people do not choose numbers at random, even when they are expected to do so²⁷. This result is in line with most studies on state lotteries. These papers show that the distribution of numbers actually chosen by players is not uniform (see, for example, Farrell *et al.* 2000, Roger and Broihanne 2007, among others) because they use common heuristics to select numbers.

The second side-result is linked to the assumption that rationality is common knowledge (Aumann 1976). It is a strong assumption and many examples show that it does not represent the way people are thinking. The most famous example is the beauty contest (first introduced by J.M. Keynes 1936, chapter 12, p. 156²⁸), translated by Moulin (1986) in numerical terms. Players have to choose a number between 0 and 100 and the winner is the one who chooses

²⁷ For example, Boland and Pawitan (1999) show that people have difficulties to choose numbers randomly, even in very simple tasks.

²⁸ As stated by John Maynard Keynes (1936): “Or, to change the metaphor slightly, professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees.”

the number closest to a given percentage (say a) of the mean choice of players. The Nash equilibrium of the game is that everybody chooses 0 when $a < 1^{29}$. However, all experiments show that most people are far from choosing 0 (Thaler 1998, Nagel 1995). Our results also confirm that a non negligible percentage of participants have a poor opinion about the rationality of others.

To sum up, we test three assumptions in this paper:

- 1) When facing positively skewed distributions, investors do not behave like risk averse expected utility maximizers. In particular, they do not use a mean-variance criterion. Positive skewness is a highly weighted decision criterion and, more precisely, the probability of the highest possible outcome is overvalued.
- 2) People are not choosing numbers at random even when they are expected to do so. They have common preferred numbers, a sub-optimal characteristic in a pari-mutuel game.
- 3) When the distribution of payoffs depends on the decision of others, agents have a tendency to consider that other people are not fully rational (and they seem right!).

The paper is organized as follows. Section 2 describes the two lottery bonds used in the survey study. This section is written in such a way that the reader can think about his/her possible answers and can “participate” to the survey. Section 3 presents the theoretical analysis of the bonds and explains what theoretical choices should be, according to the three theories under examination. Section 4 presents the empirical results and section 5 concludes.

2. Design of the lottery bonds

Lottery bonds are in general fixed-rate bonds (with coupon rate r) issued by a state or a firm. However, r applies to the global issue, not to the individual subscribers. If N one-year bonds are issued, each with a \$1 face value, the issuer repays $B = (1+r)N$ at the maturity date. A part of this amount is redistributed by means of a lottery. For example, B is divided in two parts such that $B = B_1 + B_2$. $n < N$ bonds are drawn at random and their holders share B_1 (equally or not, depending on the design of the lottery). The remaining amount, B_2 , is then shared equally among the N subscribers, or among the $N - n$ losers.

²⁹ Some guessing games involve negative feedback. Players have to find the closest number to $100 - p \times \text{mean}$. For $p = 2/3$ it is easy to see that the equilibrium choice is 60 (Sutan and Willinger 2009)

In most cases, the issuer bears no risk since it repays B whatever happens in the random draw. The risk is entirely borne by subscribers. Lottery bonds are then unusual financial assets since issuers voluntarily introduce randomness in payoffs.

In the two following subsections, we describe and characterize the two lottery bonds used in the survey study. The two bonds differ only by the way the amount paid through the lottery is defined. In the second subsection, we formalize the payoffs and introduce the notations used in the theoretical analysis of section 3.

2.1 Description of the lottery bonds

The two bonds are designed as follows.

A bank issues N (equal to 1,000,000 in the questionnaire) units of a lottery bond, each bond being sold \$1. The subscriber of one bond has to choose a number between 1 and 10. At the maturity date, the bank pays an interest rate r (5% in the questionnaire) on the global amount issued, then repaying $(1+r)N$ (\$1,050,000). However, the bank first draws one number at random between 1 and 10, say j (we say that *series j has been drawn*).

The two bonds differ in the way gains of winning subscribers are defined.

- For the first bond, the issuer pays \$1 to each of the subscribers of series j and shares equally the remaining amount among **all** subscribers, including the winning ones. For example, if $r = 5\%$, $N = 1,000$ and if 150 subscribers have chosen the winning series, they will receive \$1.90. After having paid the winners \$1 each, the bank shares the remaining \$900 among the 1,000 subscribers, each one receiving \$0.9.
- The second bond follows different rules. 10% of the initial amount issued is devoted to winners, the remaining being shared among **all** subscribers, including the winning ones. With the same data as before, a winner would receive $\$100/150 + \0.95 because the 150 winners have to share \$100. The remaining amount is constant because the global amount won through the lottery is independent of the number of winners. Consequently, all losing series receive \$0.95, whatever the number of winners is.

The main question addressed in this paper is to know how people choose a series when they get information about choices of former subscribers. Table 1 shows the individual payoffs received by subscribers of the first lottery bond, depending on the series they invested in and on the series which has been drawn. In this example, the number of bonds is 1,000,000 and the interest rate paid by the bank is 5 %, so the bank repays \$1,050,000 at the maturity date. The first line indicates the number of subscribers in each series and the first column identifies

the possible states of nature (the series number drawn at random by the issuer). There are then 10 states of nature. The following columns give the payoffs received by a subscriber of a given series in each state. For example, 1.95 is the final payoff obtained by a series-1 subscriber if number 1 is drawn. As there are 100,000 subscribers in this series, each of them first receives \$1 and the remaining \$950,000 are shared equally among all the participants, so each subscriber receives \$0.95. It explains the amounts appearing in the corresponding line. When number 2 is drawn, the series-1 subscriber receives \$0.9 because there were 150,000 subscribers in series 2. The remaining amount is \$900,000 shared by the 1,000,000 subscribers. The same calculations justify the other amounts in the table.

Table 1: Payoffs of bond 1

The “series*k*” column contains the payoffs received at the maturity date by a subscriber of series *k* when the number drawn at random is the one appearing in the first column and the same line.

	Series1	Series2	Series3	Series4	Series5	Series6	Series7	Series8	Series9	Series10
	100000	150000	80000	120000	60000	140000	70000	50000	130000	100000
1	1.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2	0.9	1.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
3	0.97	0.97	1.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
4	0.93	0.93	0.93	1.93	0.93	0.93	0.93	0.93	0.93	0.93
5	0.99	0.99	0.99	0.99	1.99	0.99	0.99	0.99	0.99	0.99
6	0.91	0.91	0.91	0.91	0.91	1.91	0.91	0.91	0.91	0.91
7	0.98	0.98	0.98	0.98	0.98	0.98	1.98	0.98	0.98	0.98
8	1	1	1	1	1	1	1	2	1	1
9	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	1.92	0.92
10	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	1.95

Table 2 shows the individual payoffs received by the subscribers of the second lottery bond, depending on the series they invested in and on the series which has been drawn. Table 2 is built as table 1, the number of subscribers in each series and the coupon rate being the same. For example, 1.95 is the final payoff obtained by a series-1 subscriber if series 1 is drawn. As there are 100,000 subscribers in this series, each of them first receives \$1 (\$100,000 shared by 100,000 winners) and the remaining \$950 000 are shared equally among all the participants. Each losing subscriber then receives \$0.95. It explains the amounts appearing in the corresponding line. When a different number is drawn, the series-1 subscriber receives \$0.95,

because the remaining amount is still \$950 000, shared among the 1,000,000 subscribers. The essential difference between the two bonds is that the payoff received by a “losing-series” doesn’t depend on the losing number, all losers receiving \$0.95. In other words, each bond 2 is characterized by only two possible payoffs, a winning one or a losing one. Only the winning amount is linked to the number of subscribers in the corresponding series.

Table 2: Payoffs of bond 2

The “series k ” column contains the payoffs received at the maturity date by a subscriber of series k when the number drawn at random is the one appearing in the first column and the same line.

	Series1	Series2	Series3	Series4	Series5	Series6	Series7	Series8	Series9	Series10
	100000	150000	80000	120000	60000	140000	70000	50000	130000	100000
1	1.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
2	0.95	1.62	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
3	0.95	0.95	2.2	0.95	0.95	0.95	0.95	0.95	0.95	0.95
4	0.95	0.95	0.95	1.78	0.95	0.95	0.95	0.95	0.95	0.95
5	0.95	0.95	0.95	0.95	2.62	0.95	0.95	0.95	0.95	0.95
6	0.95	0.95	0.95	0.95	0.95	1.66	0.95	0.95	0.95	0.95
7	0.95	0.95	0.95	0.95	0.95	0.95	2.38	0.95	0.95	0.95
8	0.95	0.95	0.95	0.95	0.95	0.95	0.95	2.95	0.95	0.95
9	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	1.72	0.95
10	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	1.95

2.2 Formal presentation of payoffs

The payoffs of bond 1 can be formalized as follows.

Let $X^i(j)$ the payoff received by a series- i subscriber when the number j is drawn. Denote N the global number of subscribers. $N = \sum_{k=1}^{10} N_k$ where N_k is the number of series- k subscribers (first line of tables 1 and 2). The issuer repays $(1+r)N$ but, according to the preceding rule, the individual payoff X^i is defined by:

$$X^i(j) = \begin{cases} 1+r - \frac{N_j}{N} & \text{if } j \neq i \\ 2+r - \frac{N_i}{N} & \text{if } j = i \end{cases} \quad (1)$$

Denote Z the random variable defined on the ten states of nature by:

$$Z(k) = (1+r)\mathbf{1}_\Omega - \theta(k), k = 1, \dots, 10 \quad (2)$$

where θ is the random variable defined on the ten states of nature by $\theta(k) = N_k / N$. $\mathbf{1}_\Omega$ is the indicator function of the set of states of nature. X^i can be written as $X^i = Z + \mathbf{1}_{\{i\}}$ with $\mathbf{1}_{\{i\}}$ being the indicator function of state i . This decomposition will be useful to study the moments of payoffs in the next section.

Suppose that you have to choose a series to invest in. Obviously, if the numbers N_j are unknown, one can reasonably expect indifference between the 10 series, which all generate an expected payoff $(1 + r)$ because the indicator functions $\mathbf{1}_{\{i\}}, i = 1, \dots, 10$, have the same probability distribution.

But what would be your choice if you were told the numbers N_j and you were the last one-unit subscriber? For example, with the data of table 1, would you choose series 8 with 50,000 subscribers, series 2 with 150,000 or a series with an intermediate number of subscribers?

The random payoffs of bond 2 can be formalized in the same way.

Let $Y^i(j)$ denote the payoff received by a series- i subscriber when the number j is drawn and still denote N the total number of subscribers. Y^i is defined by:

$$Y^i(j) = \begin{cases} 0.9 + r & \text{if } j \neq i \\ 0.9 + r + \frac{0.1N}{N_i} & \text{if } j = i \end{cases} \quad (3)$$

The random variable Y^i can be written as:

$$Y^i = (0.9 + r)\mathbf{1}_\Omega + \frac{0.1}{\theta} \mathbf{1}_{\{i\}}$$

with θ and $\mathbf{1}_{\{i\}}$ defined as before.

Suppose you have to choose a series to invest in. As before, if the numbers of subscribers N_j are unknown, one can expect indifference between the 10 series. But what would be your choice if you were told the distribution of θ and you were the last one-unit subscriber?

3. Theoretical analysis

3.1 Lottery bond 1

We remarked before that if the numbers N_j are unknown, the potential subscribers should be indifferent between series. What is changed when the information about these numbers

becomes available? The following proposition shows in particular that, if the distribution of frequencies is given, the expected return on bond 1 remains equal across series.

Proposition 1

- 1) *The expected payoff of an investment in any series of bond 1 is equal to $1+r$.*
- 2) *The variance of series- i return, conditional on a distribution $D=(N_1, \dots, N_{10})$ of bonds already sold, is decreasing with N_i and given by:*

$$V_D [X^i] = \frac{1}{10} \left[\sum_{j=1}^{10} \left(\frac{N_j}{N} \right)^2 + \left(1 - \frac{2N_i}{N} \right) \right] \quad (4)$$

Proof

- 1) Denote $E_D [X^i]$ the expected payoff received by a subscriber of series i , conditional on a given distribution D . Using the definition of Z given in equation (2), we get:

$$E_D [X^i] = E_D \left[(1+r)\mathbf{1}_\Omega - \theta + \mathbf{1}_{\{i\}} \right] = (1+r) - E_D [\theta] + E_D [\mathbf{1}_{\{i\}}]$$

The result is obtained because $E_D [\theta] = E_D [\mathbf{1}_{\{i\}}] = 0.1$. The first moment is then independent of the distribution D of frequencies. It is then not a criterion for a rational investor to decide.

- 2) Using the result of point (1) and the fact that $X^i = (1+r)\mathbf{1}_\Omega - \theta + \mathbf{1}_{\{i\}}$ we can write.

$$\begin{aligned} V_D [X^i] &= E_D \left[\left(\mathbf{1}_{\{i\}} - \theta \right)^2 \right] \\ &= E_D [\mathbf{1}_{\{i\}}] + E_D [\theta^2] - 2E_D [\theta \mathbf{1}_{\{i\}}] \\ &= 0.1 + \sum_{k=1}^{10} \left(\frac{N_k}{N} \right)^2 - 0.1 \times 2 \frac{N_i}{N} \\ &= \frac{1}{10} \left[\sum_{k=1}^{10} \left(\frac{N_k}{N} \right)^2 + \left(1 - \frac{2N_i}{N} \right) \right] \end{aligned} \quad (5)$$

Consequently, $V_D [X^i]$ is a decreasing function of N_i .

Point 2 of proposition 1 shows that a mean-variance investor would choose to “play with the crowd”, a not so intuitive result. But, if we consider the case where all subscribers choose the same number, the issue becomes a risk-free asset, paying $1+r$, whatever the number drawn by the bank. It explains why playing with the crowd is variance reducing.

3.1.1 Expected utility maximization

Consider now the general case of a risk-averse investor and denote U her utility function, assumed strictly increasing and strictly concave. The following proposition generalizes the preceding results and shows that this investor always chooses the most popular number, that is the one for which N_i is maximum.

Proposition 2

Let U denote a strictly increasing and strictly concave utility function. If $N_i < N_j$ then

$$E_D[U(X^i)] < E_D[U(X^j)]$$

Proof:

As $X^i = Z + \mathbf{1}_{\{i\}}$, assume without loss of generality that the values of Z are ranked in increasing order (corresponding to a ranking of the N_i in decreasing order). We know that $X^i(i)$ is the maximum possible value of X^i . It means that selecting number i , when buying a bond, transfers the i -th outcome of Z at the right tail of the probability distribution of X^i (since winning always generates a better outcome than losing, whatever the losing number is). It implies that transferring the lowest possible outcome to the right tail by adding \$1 is always preferred by a risk averse agent, due to the decreasing marginal utility assumption. But, as the lowest value of Z corresponds to the highest value of N_i , a risk averse investor always prefer to bet with the crowd. Proposition 2 then indicates that risk averse expected utility maximizers should choose unambiguously $n^{\circ 2}$ if they were facing the distribution of frequencies given in tables 1 and 2.

3.1.2 Optimal expectations

Optimal expectations theory is described here in a one-period framework with a finite number of states as in Brunnemeier *et al.* (2007). Agents allocate a normalized wealth of one unit to a complete set of n Arrow-Debreu securities whose prices are denoted (p_1, \dots, p_n) . The vector of subjective (objective) probabilities of states of nature is denoted Π^* (Π) and agents solve the following problem:

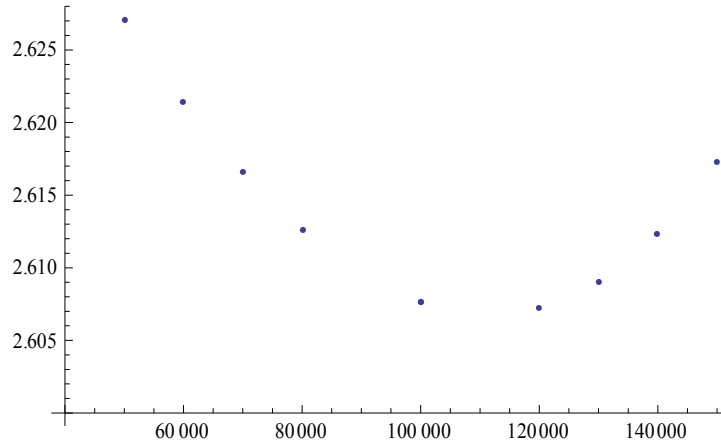
$$\max \frac{1}{2} \left(E_{\Pi^*} [U(C^*)] + E_{\Pi} [U(C^*)] \right) \quad (5)$$

under the constraint $\sum_{i=1}^n p_i c_i^* = 1$ where $C^* = (c_1^*, c_2^*, \dots, c_n^*)$. The first term $E_{\Pi^*} [U(C^*)]$ is the *ex ante* utility, C^* being the optimal choice conditional on Π^* . $E_{\Pi^*} [U(C^*)]$ may then be increased by distorting the vector of subjective beliefs Π^* . The second term in (5) is the *ex post* utility calculated with the objective probabilities. This part is penalized when C^* is far from the optimal objective choice.

Maximizing the objective function (5) means that agents choose Π^* (and consequently C^*) in order to maximize an average of the *ex ante* and *ex post* expected utilities. There is then a trade-off between an optimistic distortion of beliefs which maximizes the *ex ante* utility and the penalty coming from a sub-optimal allocation. In fact, with distorted beliefs, the optimal C^* does not lead to a maximum of the *ex post* utility $E_{\Pi} [U(C^*)]$.

Brunnemeier *et al.* (2007) get a two-fund separation result when the ratio of Arrow-Debreu security prices divided by state probabilities is constant across states of nature. The optimal portfolio consists to invest a part of wealth in the risk-free asset and the remaining amount in one and only one of the most skewed securities. In our choice context, the prices of all series are equal, the probabilities are equal and there is no aggregate risk. Therefore, the prices of Arrow-Debreu securities are identical. According to this two-fund separation result, agents should choose number 8 which is the most positively skewed portfolio. It also corresponds to the lowest N_i . Figure 1 illustrates the non monotonic link between the skewness of payoffs and the number of subscribers using the data of table 1. It appears that series 8 is the most positively skewed series. As we saw that variance is also higher for series 8, choosing this series means a strong preference for skewness and illustrates the attractiveness of the highest outcome.

Figure 1: Skewness of bond-1 payoffs for the ten series, as a function of the number of subscribers (data of table 1)



3.1.3 Cumulative prospect theory

Suppose now that agents obey CPT. They maximize a value function, depending on gains and losses, calculated with respect to a reference point. Two natural choices are available for the reference point; the initial price of the bond, or the initial price capitalized at rate r . The latter is often used when prospect theory is applied to financial decisions (Barberis *et al.* 2001, among others). The two reference points lead to the same results for the problem at hand, simply because all series' payoffs include 9 possible losses and 1 possible gain, whatever the reference point is.

In CPT, gains and losses are loaded by decision weights, obtained by distorting the cumulative (or decumulative) distribution function of payoffs. The weighting function w is defined by Tversky and Kahneman (1992) as:

$$w(p, \beta) = \frac{p^\beta}{(p^\beta + (1-p)^\beta)^{\frac{1}{\beta}}}$$

with the weighting function is different for gains and losses. For losses, decision weights are obtained by applying w to the cumulative distribution function. For gains, it is applied to the decumulative distribution function. Moreover, the parameter β may be different for gains and losses. The values estimated by Tversky and Kahneman (1992) were $\beta^+ = 0.61$ and $\beta^- = 0.69$. In our framework, all states have the same probability 0.1. When outcomes are ranked in increasing order, the weight of the k -th outcome is then $\pi_k = w[0.1 \times k, \beta^-] - w[0.1 \times (k-1), \beta^-]$. The weight of the unique gain is equal to $w(0.1, \beta^+)$. We observe here that weights are determined only by the ranking of outcomes, not by their values. It is also important to notice that the unique gain would allow to use a unique weighting function, as in rank-dependent utility models. In this framework, the weight of the

gain is $1 - w(0.9, \beta^-)$ and the decision weights define a probability measure on the set of states of nature. This case will be examined hereafter.

The value function is defined, for a gain/loss x by:

$$v(x) = \begin{cases} x^\alpha & \text{if } x > 0 \\ -\lambda(-x)^\alpha & \text{elsewhere} \end{cases}$$

$\lambda > 1$ is the loss aversion coefficient and $0 < \alpha < 1$ characterizes the curvature of the value function. v is then concave for gains and convex for losses. Investors then choose number i to maximize $\sum_{k=1}^{10} \pi_k v(X^i(k) - x^*)$ where x^* is the reference point.

To understand what would be the choice of a CPT investor, we report on table 3 some elements of comparison for the two series with the lowest number of former subscribers, $N_i = 50,000$ (series 8) and $N_j = 60,000$ (series 5). We observe that the two cumulative distribution functions are equal for the 8 lowest payoffs, as shown in the two first lines of table 3.

Table 3: Payoffs of series 5 and 8

The payoffs of the two series are ranked in increasing order and the corresponding weights are calculated with the parameters estimated by Tversky and Kahneman (1992), except β which is equal for gains and losses. We then have $\alpha = 0.88$; $\lambda = 2.25$; $\beta = 0.65$.

Series5	0.9	0.91	0.92	0.93	0.95	0.95	0.97	0.98	1	1.99
Series8	0.9	0.91	0.92	0.93	0.95	0.95	0.97	0.98	0.99	2
Weight	0.17872	0.08121	0.06432	0.05798	0.05872	0.05654	0.06493	0.07759	0.10545	0.17872

The difference between the two series comes from the two highest payoffs which are (1;1.99) for series 5 and (0.99; 2) for series 8. The third line gives the decision weights for $\beta = 0.65$. The ratio of weights for the two highest payoffs is around 1.78. It is then an incentive to choose series 8 with the highest payoff. However, two other variables are important in the choice. The loss aversion coefficient, equal to 2.25 in the table, favors series 5 and largely compensates the lower weight. Moreover, the gain in the state with the highest payoff (0.94 or 0.95 with a reference point of 1.05) is much higher than the absolute value of the loss in the other state (0.05 or 0.06). As $\alpha < 1$, the increase of the value function between 0.05 and 0.06 is much larger than the absolute value of the decrease between 0.95 and 0.94. The consequence is that a CPT agent would prefer series 5 instead of series 8 for a large range of

parameters. Moreover, if α is decreased, then increasing the curvature of the value function, the optimal CPT choice goes to a series with a larger number of subscribers. If the loss aversion coefficient is decreased, the result is unchanged (the optimal choice is not series 8) as long as loss aversion stays above 1.2, an unusual value in experimental studies. Figure 2 shows the CPT evaluation of the 10 series with the initial parameters. Using a weighting function only based on the distortion of cumulative distribution functions, as in the rank-dependent expected utility model, induces a weight of 0.25 for the gain which favors series n°8. However, it is not enough to compensate loss aversion and the effect of decreasing marginal utility. Series 5 with 60000 subscribers is the optimal choice in this case, as long as the loss aversion coefficient is greater than 1.72.

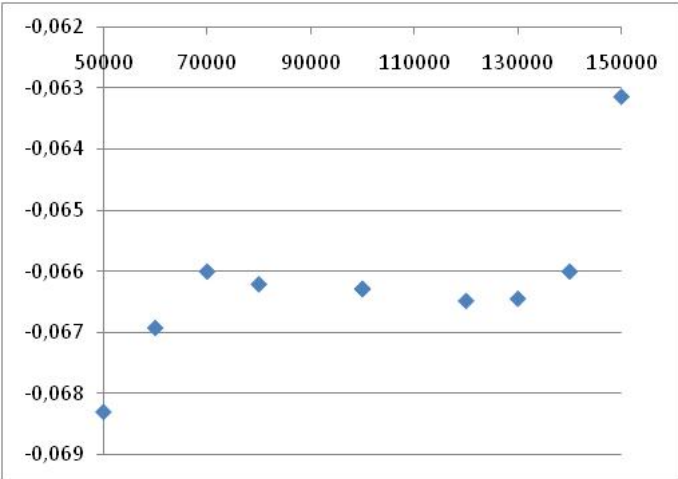


Figure 2: Value function of CPT for each series.

The horizontal axis gives the number of subscribers and the vertical axis the value function of CPT with $r = 1.05$. The reference point is $r = 1.05$.

To summarize the analysis (using the data of table 1), agents should choose series 2 if they are risk averse expected utility maximizers, series 8 if their behavior is driven by optimal expectations theory and, finally, series 2 or 5 if they obey cumulative prospect theory with usual parameter values. For CPT, changing the parameters could lead to other choices but keeping reasonable values leads to a choice different from series 8. This result may seem surprising since CPT is often used to justify the participation to unfair state lotteries. However, in state lotteries, the probability of winning is very low, giving a more important role to the distortion of beliefs introduced in CPT. Here, the objective probability of the

winning state is only multiplied by 1.78. It is not enough to make the choice of the most skewed alternative optimal.

3.2 Lottery bond 2

The analysis of bond 2 is much more simple. We saw in equation (3) that:

$$Y^i(j) = (0.9+r)\mathbf{1}_\Omega + \frac{0.1N}{N_i}\mathbf{1}_{\{i\}} = \begin{cases} 0.9+r & \text{if } j \neq i \\ 0.9+r + \frac{0.1N}{N_i} & \text{if } j = i \end{cases}$$

Losing series generate the same payoff $0.9+r$ but the payoff of the winning series is inversely proportional to the number of subscribers in this series. Being given a distribution of frequencies, the optimal choice of the last subscriber of the issue is always the series with the lowest number of former subscribers. In fact, the corresponding payoff dominates the others for first order stochastic dominance. Looking more closely at two series i and j with frequencies N_i and N_j with $N_i < N_j$, we observe that the payoffs are $0.9+r$ with probability

0.9 but series i pays $0.1 \times N \times \left(\frac{1}{N_i} - \frac{1}{N_j} \right)$ more than series j with probability 0.1 . It is as if you

were given for free a lottery ticket paying this amount with probability 0.1 . Whatever your preferences are (obeying first order stochastic dominance), you accept the lottery ticket. Bond 2 is used in the survey study to introduce a screening process. Due to the problem addressed in the paper, there is no incentive compatible payment scheme because we have no a priori about which choice is the “right” one for bond 1. However, we can suspect that the answers of participants which do not obey first-order stochastic dominance are highly questionable. Simply, some students may not be motivated by the exercise and may answer at random. The empirical section is then focused on participants having provided answers compatible with first order stochastic dominance. Nevertheless, we provide in the appendix (table A3) the answers for the complete sample.

3.3 Rationality as common knowledge

Assume now that you have to choose a series to invest in bond 2, after one million other subscribers but also assume that one more million subscribers will choose after you, with updated information about sales (here 2,000,000 bonds have been issued). If rationality is common knowledge, it is not difficult to see that the equilibrium sharing of bonds at the end of the choice process should be an equal sharing across series. In fact, for bond 2, it is always optimal to choose the series with the lowest frequency when you are the last subscriber. It

implies that, as long as the frequency of a given series is lower than 200,000, you can choose this series. The following rational subscribers will stop choosing a given series when the frequency will reach 200,000. Beyond this threshold, this choice becomes sub-optimal because there is at least another series with a lower number of subscribers. Consequently, if you believe that rationality is common knowledge, you can choose at random if the current sharing of bonds is the one given in table 1 or 2.

For bond 1, the story is a little bit different. If you think that others are like you, you should choose the same answer to questions 3 and 5 (see the questionnaire in the appendix) if you are a risk-averse expected utility maximizer, assuming that the following subscribers will also bet with the crowd. The optimal choice if you obey OET is, as for bond 2, to invest at random, assuming the others also obey OET. Obviously, if you consider that a proportion of agents is risk-averse, you will never choose the highest frequency series, anticipating that it will be chosen by these expected utility maximizers.

3.4 The random choice of numbers

Finally assume that you have no information about the sharing of bonds across former subscribers. You are only told the way the bank will reimburse the issue. In this case, the probability distribution of returns is equal in each series, either for bond 1 or for bond 2. Your choice then should be random in the set of ten series. Therefore, the distribution of choices at the aggregate level should be uniform. We show in the next section that it is not the case.

4. The survey study

4.1 The questionnaire

The study was realized during different finance courses in two French universities (University of Strasbourg and University of Clermont-Ferrand) and two business schools, EM Strasbourg Business School (France) and HEC Lausanne (Switzerland). 337 students participated, enrolled in economics, finance or accounting programs at the MSc level³⁰. The complete questionnaire is provided in the appendix. The English version was used in Lausanne (all courses being taught in English at the MSc level) and a French version in the other programs. The numbers of students in the different locations are provided in table 4.

Table 4: Origin of participants

³⁰ For the University of Strasbourg, 86 (40) students come from the MSc in Finance (Actuarial Studies), in Lausanne, 69 (5) come from the MSc in Finance (Actuarial Studies), in Clermont Ferrand, there were 43 students from the MSc in Finance, 39 from the MSc in Accounting and Control and 14 from the MSc in Economic Analysis.

University or Business School	Number of students
EM Strasbourg Business School	41
University of Strasbourg	126
HEC Lausanne	74
University of Clermont-Ferrand	96
TOTAL	337

The participants had to answer 6 questions, divided into three groups of 2 questions, related to the lottery bonds presented in the preceding section. In each pair of questions, the first one is related to bond 1 and the second to bond 2. The required answers were simply numbers between 1 and 10 corresponding to the choice of a series number.

For the first two questions, participants were only told the characteristics of the bonds, without any other information, either on the table of payoffs or on the choice of former subscribers. For example, bond 1 was presented as follows.

*A bank (called bank1) issues 1,000,000 one-year bonds at a price of 1€ each. The bank repays 1,050,000€ at the end of the year (a 5% interest rate on the issue). When buying one bond, subscribers choose an integer number between 1 and 10. On the repayment date, the bank draws at random a (lucky) number between 1 and 10 and first repays 1€ to each subscriber having chosen the lucky number. The remaining amount is equally shared among **all** subscribers, including the winning ones.*

As mentioned before, with no information other than the way payoffs are defined, investors should be indifferent between the series; we then expect a random choice for the two first questions.

After having answered the two first questions, participants received the information summarized in table 5 (it corresponds to the first line of tables 1 and 2). It was provided on a slide, so all participants knew that everybody was receiving the same information. For questions 3 and 4, it was specified that the respondent was about to buy the last bond of the issue. In other words, everybody was able to infer the final distribution of payoffs.

Table 5: Information about choices of former subscribers (shown publicly on a Powerpoint slide)

Quantities already bought in each series				
Series1	Series2	Series3	Series4	Series5
100 000	150 000	80 000	120 000	60 000
Series6	Series7	Series8	Series9	Series10
140 000	70 000	50 000	130 000	100 000

A slight modification was introduced for the 96 students at university of Clermont-Ferrand. They were shown table 5 for questions related to bond 1 (questions 3 and 5) and table A2 (see the appendix) for questions 4 and 6. In table A2, the quantities of bonds are the same but the numbers associated with these quantities are different. We then deal with the same set of cumulative distributions of returns but the numbers identifying a given distribution are not the same. The idea was to control for a possibility of “inertia” in the answers. We saw, in the preceding section, that the rational answer for question 4 is series n°8 but, it is also the answer to question 3 for agents obeying optimal expectations theory. Consequently, we had to check if some students were choosing the same number to answer the two questions, simply by applying a law of least effort. In fact, nobody in the Clermont-Ferrand subsample (96 students) chose (8, 8) to answer questions 3 and 4. There is then no reason to think that it is different for the other subsamples. In the subsample of 96 students, the rational answer to question 4 was n°4. We then also checked whether some students could have been considered rational “by inertia”, that is by selecting the same answer for questions 3 and 4. More precisely, we counted the number of students having chosen the same number for questions 3 and 4, whatever this number was. Only 3 students made such a choice over the subsample of 96 students. Consequently, the results presented in the following cannot be invalidated with this argument.

In the third sequence of two questions (questions 5 and 6), the rule was that participants had to choose a series number with the same information as in questions 3 and 4, but they were told that one million bonds were still to be sold to other subscribers after their own choice. Moreover, participants were also informed that the future subscribers would get updated information about sales at the time of their own purchase. Therefore, participants had to build expectations about the decision rules of future subscribers. These two last questions are devoted to analyze the opinion of participants about the rationality of others, as in the usual beauty contest.

To present the results concerning questions 3 to 6 in a simple way, we use the numbering of series in table 5 for all the subsamples. Obviously, for questions 1 and 2, participants were not shown table 5 or table A2, so we keep the numbers they used to answer.

4.2 Results

4.2.1 Attractiveness of the best outcome

As mentioned before, the data provided in table 5 imply that agents obeying first-order stochastic dominance must choose series n°8 at question 4. 245 students over 337 made this choice. They are called “rational” in the following even if “wrong” answers to this question can simply be due to a lack of motivation to participate. Table 6 shows the percentage of rational answers in each training program.

The lower percentage in Clermont-Ferrand is possibly due to the more complicated task students had to manage with different data for bonds 1 and 2. Apart from this, the results are not really surprising. We find the highest percentages in the two Business School programs in Strasbourg and Lausanne. The other programs are part of Departments of Economics and Management Science. University students are less used (and possibly less motivated) to participate in such surveys than students in Business Schools³¹. It may explain the significant difference between the proportions.

Table 6: Numbers of « rational » participants across training programs

	Rational answers	Total	Percentage "rational"
HEC Lausanne	60	74	81,08%
EM Strasbourg Business School	34	41	82,93%
University of Strasbourg	90	126	71,43%
University of Clermont-Ferrand	61	96	63,54%
Total	245	337	72,70%

The detailed answers to the other 5 questions are given in table 7³², panel A. The lines of the table are ranked according to the number of hypothetical subscribers in the series, not according to the series number. The main point concerns question 3. There is a clear

³¹ For example, students in Business Schools are used to evaluate courses and teachers, a not so common practice in French universities.

³² Table A3 in the appendix gives (for completeness of information) the answers of the global sample, including the participants qualified as “irrational”.

preference for series n°8 since 56.61% of respondents chose this series. The second important observation is the 22.31% of participants having chosen series 2. They are either risk averse expected utility maximizers or CPT investors. The eight other possible answers have almost negligible frequencies since they gather around 20%. However, as mentioned in section 3, these answers could possibly be attributed to CPT investors with different parameters or weighting functions.

Our results show a strong preference, not only for skewed returns (all returns are skewed in our study) but for the most skewed return, and for the random payoff with the highest outcome. This result is then clearly in line with the predictions of optimal expectations theory.

It is also interesting to come back to the comparison of cumulative distribution functions of series 8 and 5 (provided in table 3), corresponding to the two first lines of table 7. It is remarkable that 56.61% of participants chose series 8 and only 2.89% series 5 when the difference between the two is only a swap of one cent between the two highest outcomes. It confirms the attractiveness of the highest outcome, as predicted in OET.

Panel B of table 7 provides the proportions of choices 2 and 8 in the different locations. The proportions in the two universities are very close to each other. No significant difference appears in the distributions. On the contrary, we observe a higher (lower) proportion of expected utility maximizers in the sample of HEC Lausanne (EM Strasbourg) but it is difficult to interpret these differences, taking into account the number of students in each subsample.

Table 7

Panel A: Answers of « rational » participants

The figures are provided in percentage but the number of answers in each column varies from 241 and 244 (a few students left some questions unanswered). The answers for Q4 are not provided because, by construction, all “rational” respondents chose series n°8 for this question (except those in Clermont-Ferrand for which the rational answer was n°4).

BONDS BOUGHT	Q1	Q2	Q3	Q5	Q6
50000	8,30	8,26	56,61	20,25	27,05
60000	12,03	13,22	2,89	4,13	4,10
70000	16,60	15,29	3,72	8,68	9,02
80000	10,37	9,09	4,13	7,85	5,74
100000	15,77	15,29	2,89	6,61	6,15
100000	7,47	9,09	1,24	2,48	5,74
120000	7,05	5,37	1,24	2,48	2,46
130000	7,88	7,44	2,07	2,48	1,64
140000	3,73	4,96	2,89	6,61	7,38
150000	10,79	11,98	22,31	38,43	30,74
TOTAL	100,00	100,00	100,00	100,00	100,00

Panel B: Percentages of choices 2 and 8 in the different locations

EMS= EM Strasbourg Business School, LAU = HEC Lausanne, UDS = University of Strasbourg, UCF = University of Clermont-Ferrand. N is the number of “rational” participants in each subsample. The last line gathers all answers that do not correspond to choices 2 and 8.

	EMS	LAU	UDS	UCF
N	34	60	90	61
Series 8	82,35%	46,67%	54,44%	52,46%
Series 2	5,88%	31,67%	23,33%	19,67%
Other	11,76%	21,67%	22,22%	27,87%

4.2.2 Opinions about the rationality of others

Questions 5 and 6 were designed as questions 3 and 4, except that participants were told that one million bonds were still to be sold after their own choice, the next subscribers choosing with updated information about sales.

As mentioned in the preceding section, concerning Q6, if participants were thinking that other subscribers are rational they should be indifferent between all solutions. We then expect a uniform distribution of choice. However, assuming that other subscribers are not completely rational, and have a “one-step” reasoning, leads you to play with the crowd, expecting that the others will stay in the low frequency series. In the same way, with a two-step reasoning, you should stay in the low frequency series. Figure 3 shows that answers to question 6 corresponds to agents mainly using a one or two-step reasoning. 141 participants (58%) chose

numbers 2 or 8 corresponding to the two highest frequencies. Obviously, the uniform distribution hypothesis is rejected at conventional levels.

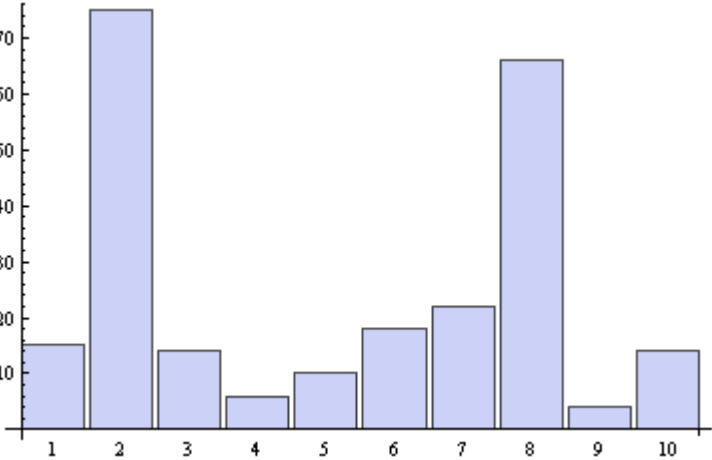


Figure 3: Bar chart of answers to question 6

Concerning questions 3 and 5, the panel A of table 8 summarizes the frequencies of choice of the 242 participants having answered the two questions. The two series with 100,000 former subscribers have been aggregated since they pay exactly the same payoffs. The first column (row) gives the answer to question 3 (5). We observe that among the 54 participants having chosen the highest frequency at question 3, 39 (72.2%) stay on that choice for question 5. It is rational since they want to bet with the crowd and hope that the others will do the same. We also observe the behavior already mentioned for question 6. Around one third of participants (47) obeying OET use a one step reasoning and switch to the highest frequency series at question 5. They expect that the others will continue to bet on the lowest frequency series (which then becomes a high frequency series!). 35 participants (25%) stay on their choice, using a two-step reasoning. The remaining 40% have a more sophisticated behavior by choosing other series. One more time, it is clear that the conditional distribution of answers to question 5 (conditioned on the choice of the lowest frequency series to question 3) is not uniform. The deviation from the uniform distribution ($\chi^2 > 140$)³³ is essentially due to the participants using a one-step or a two-step reasoning. They account for more than 75% of the χ^2 value.

³³ To calculate the χ^2 statistic, we grouped the two series with 120,000 and 130,000 subscribers due to the low frequencies of these series.

If we compare now the answers to questions 3 and 6 (table 8, panel B), almost the same comments can be done. We also reject the uniform distribution assumption for question 6, conditional on the choice of the lowest frequency at question 3 ($\chi^2 \approx 116$).

Table 8: Answers to the pairs of questions (3, 5) and (3, 6)

Panel A: Frequencies of pairs of choices for questions 3 and 5

Q3\5	50000	60000	70000	80000	100000	120000	130000	140000	150000
50000	35	5	14	7	13	4	3	9	47
60000	1	2	0	0	0	0	0	1	3
70000	0	0	7	1	1	0	0	0	0
80000	1	0	0	6	1	0	0	1	1
100000	4	0	0	1	5	0	0	0	0
120000	1	0	0	0	0	2	0	0	0
130000	0	0	0	0	0	0	2	0	3
140000	1	0	0	1	0	0	1	4	0
150000	6	3	0	3	2	0	0	1	39

Panel B: Frequencies of pairs of choices for questions 3 and 6

Q3\6	50000	60000	70000	80000	100000	120000	130000	140000	150000
50000	30	6	13	8	14	3	3	12	48
60000	1	0	0	1	0	0	0	1	4
70000	3	0	4	0	1	0	0	1	0
80000	3	1	1	1	0	0	0	1	3
100000	5	0	0	1	1	0	0	1	2
120000	0	0	1	0	0	0	0	0	2
130000	4	0	1	0	0	0	0	0	0
140000	2	0	0	0	2	2	0	1	0
150000	17	3	2	3	11	1	1	1	15

To sum up, it appears that the choices of participants do not confirm the assumption that rationality is common knowledge. The most frequent answers correspond to a one-step or two-step reasoning. This result is indeed not surprising since it is the most common values found in the literature for the depth of reasoning³⁴.

4.2.3 Heuristics in random choice of numbers

We expect random choices to questions 1 and 2 because participants have no information about choices of former subscribers. Figure 4 shows the bar chart cumulating choices for these two questions. In case of a uniform distribution, we should have frequencies around

³⁴ See Nagel (1999) for a survey on beauty contest games.

48.3 which is the mean frequency of the uniform distribution. It appears that the actual answers are far from a uniform distribution. This hypothesis is clearly rejected at the 1% level. The χ^2 value is 63.97 for a 1% critical value of 21.66.

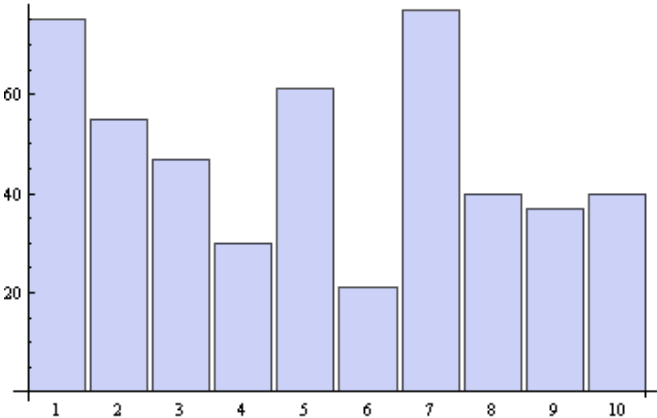


Figure 4: Bar chart of answers to questions 1 and 2

The popularity of numbers 7 and 5, representing respectively 15.94% and 12.62% of choices for questions 1 and 2, can be explained by comparable results for state lotteries. For example, in the French lotto game, players have to choose 5 numbers between 1 and 49 and (independently) a lucky number between 1 and 10. The sponsor of the game draws at random the winning combination and the lucky number. Since the start of this version of the game in October 2008, number 7 has been drawn 14 times³⁵ (see the detailed results in table 9). For these particular draws, the mean proportion of winners of the lucky number is 16.53% when 10% is expected if players choose their numbers at random³⁶. For number 5, the mean proportion is 13% in a set of 10 draws, the minimum and maximum proportions being 12.16% and 13.79%.

The main difference between the results of the lotto game and ours concerns number 1. However, this difference can possibly be justified by the fact that participants, considering all choices as equivalent, select number 1 which is obviously the first in the list of possible choices. In some sense, selecting number 1 could be interpreted as a random choice. But even if we share these answers between the ten numbers, the difference with a uniform distribution remains significant.

Table 9: Relative frequencies of numbers chosen by French lotto players

³⁵ For the 155 first draws up to 09/26/2009.

³⁶ The data on French lotto draws are provided on www.fdjeux.com and the percentage of lucky number winners is reported on www.sojah.com.

The column “Mean” gives the proportion of players having chosen the number in the first column over the number of draws in the fifth column. “Min” and “Max” are the corresponding minimum and maximum proportion over the same number of draws.

Number	Mean	Min	Max	Number of draws	Standard deviation
1	7,85%	7,36%	9,29%	17	0,56%
2	8,29%	7,77%	9,40%	14	0,38%
3	11,30%	10,76%	11,68%	18	0,22%
4	10,67%	9,72%	11,28%	20	0,35%
5	13,00%	12,16%	13,79%	10	0,51%
6	10,20%	9,82%	10,77%	14	0,27%
7	16,53%	15,75%	17,10%	14	0,37%
8	10,27%	9,92%	10,75%	17	0,22%
9	9,41%	8,94%	9,86%	15	0,22%
10	6,46%	5,98%	6,94%	16	0,26%

5. Conclusion

In this paper, we analyze the way people manage a simple financial decision making problem based on lottery bonds. Our purpose was to compare three theories, expected utility, cumulative prospect theory, and the more recent optimal expectations theory. The latter implies that agents are especially attracted by the best outcome and overvalue its probability of occurrence. To distinguish between skewness seeking and attractiveness of the best outcomes we designed a lottery bond in such a way that all choices generate positively skewed distributions. We show that more than 55% of participants to the survey select the choice with the highest outcome, controlling for the expected return. Moreover, this choice is dominated by all other choices when expected utility is used as the decision making model. Our results show that gambling and investing cannot be treated separately, an idea that appeared in several recent papers. These results are then consistent with the theoretical analysis of Brunneimeier *et al.* (2007) and with the empirical study of Bali *et al.* (2009).

We also showed that people use heuristics to choose numbers at random, leading to non random choices at the aggregate level. Numbers 7 and 5 are especially popular and it comes

with no surprise³⁷. Finally, by introducing a kind of beauty contest in the questionnaire, we observe that respondents do not assume that rationality is common knowledge, either because they recognize their own limited rationality or because they consider that others are not fully rational. The questions used in this paper are very simple and make the study easily replicable. We then hope that it will be replicated on other populations with different cultural backgrounds to test if our results can be generalized to alternative environments.

³⁷ For example, Roger and Broihanne (2007) show that these numbers are among the most popular in the French lotto game.

References

- Alpert, M., H. Raiffa. 1982. A Progress Report on the Training of Probability Assessors, in D. Kahneman, P. Slovic and A. Tversky, eds, *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, 294-305.
- Aumann, R. 1976. Agreeing to Disagree, *Annals of Statistics* 4(6), 1236–1239.
- Bali, T.G., N. Cakici, R.F. Whitelaw. 2009. Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns, NBER Working Papers 14804.
- Barberis, N. M. Huang. 2008. Stocks as Lotteries: The Implications of Probability Weighting for Security Prices, *American Economic Review*, 98(5), 2066-2100.
- Barberis, N., M. Huang, T. Santos. 2001. Prospect Theory and Asset Prices, *Quarterly Journal of Economics*, 116, 1-53.
- Boland, P.J., Y. Pawitan. 1999. Trying to be random in selecting numbers for Lotto, *Journal of Statistics Education*, 7(3).
- Bosch, A., J. Garcia-Montalvo, R. Nagel, A. Satorra. 2002. One, Two, (Three), Infinity,...: Newspapers and Lab Beauty Contest Games, *American Economic Review*, 92(5), 1687-1701.
- Brunnermeier, M.K., C. Gollier, J.A. Parker. 2007. Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns, *American Economic Review*, 97(2), 159-165.
- Brunnermeier, M.K., J.A. Parker. 2005. Optimal Expectations, *American Economic Review*, 95(4), 1092-1118.
- Buehler, R., D. Griffin, M. Ross. 1994. Exploring the “Planning Fallacy”: Why People Underestimate Their Task Completion Times, *Journal of Personality and Social Psychology*, 67(3), 366-381.
- Calvet, L., J.Y. Campbell, P. Sodini. 2007. Down or Out: Assessing the Welfare Costs of Household Investment Mistakes, *Journal of Political Economy*, 115, 707-747.
- Farrell, L., R. Hartley, G. Lanot, I. Walker. 2000. The Demand for Lotto: the Role of Conscious Selection, *Journal of Business & Economic Statistics*, 18(2), 228–241.
- Goetzmann, W. N., A. Kumar. 2008. Equity Portfolio Diversification, *Review of Finance*, 12(3), 433-463.
- Green, R.C., K. Rydqvist. 1997. The Valuation of Non-Systematic Risks and the Pricing of Swedish Lottery Bonds, *Review of Financial Studies*, 10, 447-480.

- Guillen, M.F., A.E. Tschoegl. 2002. Banking on Gambling: Banks and Lottery-Linked Deposit Accounts, *Journal of Financial Services Research*, 21(3), 219-231.
- Harvey, C., A. Siddique. 2000. Conditional skewness in asset pricing tests, *Journal of Finance*, 55, 1263-1295.
- Kahneman, D., A. Tversky. 1979. Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47, 263-291.
- Keynes, J. M. 1936. *The General Theory of Employment, Interest and Money*, Harcourt Brace and Co.
- Kraus, A., R.H. Litzenberger. 1976. Skewness preference and the valuation of risk assets, *Journal of Finance*, 31, 1085-1100.
- Kumar, A. 2009. Who Gambles on the Stock Market?, *Journal of Finance*, 64(4), 1889-1933.
- Lévy-Ullmann, H. 1896. Lottery Bonds in France and in the Principal Countries of Europe, *Harvard Law Review*, 9, 386-405.
- Millar, J.R., J.A. Gentry 1980. The Soviet Experiment with Domestic Lottery Bonds, *Financial Management*, Winter, 21-29.
- Mitton, T., K. Vorkink. 2007. Equilibrium Underdiversification and the Preference for Skewness, *The Review of Financial Studies*, 20(4), 1255-1288.
- Moulin, H. 1986. *Game Theory for the Social Sciences*, 2nd Ed., NYU Press.
- Nagel, R. 1995. Unraveling in Guessing Games: An Experimental Study, *American Economic Review*, 85, 1313-1326.
- Nagel, R. 1999. A Survey of Experimental Beauty Contest Games : Bounded Rationality and Learning, *Games and Human Behaviour, Essays in Honor of Amnon Rapoport*, eds. David V. Budescu, Ido Erev and Rami Zwick. Lawrence Erlbaum Associates, Inc. New Jersey, 105-142.
- Pfiffelmann, M., P. Roger. 2005. Les comptes d'épargne associés à des loteries : approche comportementale et étude de cas, *Bankers, Markets and Investors*, 78, September-October, 1-8.
- Polkovnichenko, V. 2005. Household Portfolio Diversification: A Case for Rank-Dependent Preferences, *The Review of Financial Studies*, 18(4), 1467-1501.
- Quiggin, J. 1982. A Theory of Anticipated Utility, *Journal of Economic Behavior & Organization*, 8, 641-645.
- Ridge, J., M. Young. 1998. Innovations in Savings Schemes: The Bonus Bonds Trust in New Zealand, *Financial Services Review*, 7(2), 73-81.

Roger, P., M.H. Broihanne. 2007. Efficiency of Betting Markets and Rationality of Players: Evidence from the French 6/49 Lotto, *Journal of Applied Statistics*, 34(6), 645-662.

Sutan, A., M. Willinger. 2009. Guessing with Negative Feedback: An Experiment, *Journal of Economic Dynamics and Control*, 33, 1123-1133.

Thaler, R. 1998. Giving Markets a Human Dimension, in *The Complete Finance Companion*, Financial Times/Prentice Hall.

Tufano, P. 2008. Saving whilst Gambling: An Empirical Analysis of UK Premium Bonds, *American Economic Review*, 98(2), 321-326.

Tversky, A., D. Kahneman. 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, 12, 297-323.

Weinstein, N.D. 1980. Unrealistic Optimism About Future Life Events, *Journal of Personality and Social Psychology*, 39(5), 806-820.

APPENDIX: Questionnaire

BOND n°1

A bank (called bank1) issues 1,000,000 one-year bonds at a price of 1 € each. The bank repays 1,050,000 € at the end of the year (a 5% interest rate on the issue). When buying one bond, subscribers choose an integer number between 1 and 10. On the repayment date, the bank draws at random a (lucky) number between 1 and 10 and first repays 1 € to each subscriber having chosen the lucky number. The remaining amount is equally shared among **all** subscribers, including the winning ones.

- **Q1: Suppose that you buy one bond, which number do you choose?**

BOND n°2

A bank (called bank2) issues 1,000,000 one-year bonds at a price of 1 € each. The bank repays 1,050,000 € at the end of the year (a 5% interest rate on the issue). When buying one bond, subscribers choose an integer number between 1 and 10. On the repayment date, the bank draws at random a (lucky) number between 1 and 10 and first shares equally 100,000 € among subscribers having chosen the lucky number. The remaining amount is equally shared among **all** subscribers, including the winning ones.

- **Q2: Suppose that you buy one bond, which number do you choose?**

The table on the screen gives the number of bonds already sold for each choice of the lucky number.

- **Q3:** If you are about to buy the **last** bond issued by bank 1, which number do you choose?
- **Q4:** If you are about to buy the **last** bond issued by bank 2, which number do you choose?
- **Q5:** Assume now that 2,000,000 bonds were issued by bank 1. If you buy a bond, knowing that 1,000,000 more bonds are to be sold after your choice (the next subscribers being fully informed about the evolution of choices by the updating of the table on the screen), which number do you choose?
- **Q6:** Assume now that 2,000,000 bonds were issued by bank 2. If you buy a bond, knowing that 1,000,000 more bonds are to be sold after your choice (the next subscribers being fully informed about the evolution of choices by the updating of the table on the screen), which number do you choose?

Table A1

Table shown (on a Powerpoint slide) to the participants after they answered questions 1 and 2

Quantities already bought in each series				
Series1	Series2	Series3	Series4	Series5
100 000	150 000	80 000	120 000	60 000
Series6	Series7	Series8	Series9	Series10
140 000	70 000	50 000	130 000	100 000

Table A2

Table shown (on a Powerpoint slide) to the 96 participants in Clermont Ferrand for questions 4 and 6

Quantities already bought				
Series1	Series2	Series3	Series4	Series5
140 000	70 000	130 000	50 000	100 000
Series6	Series7	Series8	Series9	Series10
100 000	120 000	80 000	150 000	60 000

Table A3

Results for the complete sample of 337 participants

SERIES NUMBER	BONDS BOUGHT	Q1	Q2	Q3	Q4	Q5	Q6
1	100000	56	58	14	10	30	25
2	150000	33	35	75	41	116	106
3	80000	38	33	19	8	24	21
4	120000	24	24	8	5	11	10
5	60000	48	45	16	12	20	15
6	140000	13	14	13	9	17	21
7	70000	49	46	12	4	25	26
8	50000	25	25	160	245	70	84
9	130000	21	19	9	1	8	8
10	100000	26	33	8	2	12	19